

WEST Search History

[Hide Items](#)[Restore](#)[Clear](#)[Cancel](#)

DATE: Wednesday, January 14, 2004

Hide?	Set Name	Query	Hit Count
		<i>DB=PGPB,USPT; PLUR=YES; OP=ADJ</i>	
<input type="checkbox"/>	L10	l5 and insect resistanc\$	2
<input type="checkbox"/>	L9	l5 and insect resistance	0
<input type="checkbox"/>	L8	L5 and transgenic	39
<input type="checkbox"/>	L7	L5 and insect attack	2
<input type="checkbox"/>	L6	L5 and control	48
<input type="checkbox"/>	L5	l1 and plant	48
<input type="checkbox"/>	L4	L1 and (coat protein precursor or capsid precursor)	3
<input type="checkbox"/>	L3	L1 and (rna dependent rna polymerase or rdrp)	2
<input type="checkbox"/>	L2	L1 and (replicase or coat protein)	19
<input type="checkbox"/>	L1	hasv or helicoverpa armigera stunt virus	91

END OF SEARCH HISTORY

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 17:34:30 ; Search time 8964 Seconds
(without alignments)
11309.020 Million cell updates/sec

Title: US-09-991-262-47
Perfect score: 2478
Sequence: 1 GTTTTCTTTCTTTACCAAG.....CTGCGTCCTCGGATACCA 2478

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2478	100.0	2478	6	ARI25096	ARI25096 Sequence
2	2478	100.0	2478	6	ARI25097	ARI25097 Sequence
3	2478	100.0	2478	6	ARI59968	ARI59968 Sequence
4	2478	100.0	2478	6	AX028950	AX028950 Sequence
5	2478	100.0	2478	6	AX028971	AX028971 Sequence
C	2478	100.0	2478	14	HVSC0AT	L37299 Helicoverpa
7	2467	99.6	2479	6	ARI25098	ARI25098 Sequence
8	833	33.6	2448	14	S43937	S43937 capsid prot
9	144.8	5.8	2593	14	AF548354	AF548354 Providenc
10	71.2	2.9	5312	6	ARI25092	ARI25092 Sequence
11	71.2	2.9	5312	6	ARI25093	ARI25093 Sequence
12	71.2	2.9	5312	6	ARI25094	ARI25094 Sequence
13	71.2	2.9	5312	14	HAI18246	U18246 Helicoverpa
14	71.2	2.9	5368	6	ARI25095	ARI25095 Sequence
C	62	2.5	125020	9	AF429315	AF429315 Homo sapi
15	61.6	2.5	2000	6	AX655393	AX655393 Sequence
C	60	2.4	6534	6	ARI59969	ARI59969 Sequence
17	60	2.4	6536	6	AX028953	AX028953 Sequence
C	60	2.4	6536	6	AX028972	AX028972 Sequence
19	58	2.3	125020	9	AF429315	AF429315 Homo sapi
20	57.2	2.3	3030	8	VCA429230	AJ429230 Volvox ca
21	55.6	2.2	2000	6	AX655393	AX655393 Sequence
C	54.2	2.2	310550	1	SCO939113	AL939113 Streptomy
23	54	2.2	1910	6	AX653865	AX653865 Sequence
24	54	2.2	126962	2	AC109595	AC109595 Oryza sat
C	52	2.1	154138	2	AC129716	AC129716 Oryza sat
26	52	2.1	327650	1	EX248337	EX248337 Mycobacte
27	50.2	2.0	178827	8	CNS08C89	AL732381 Oryza sat
C	50.2	2.0	193499	8	CNS08C94	AB014924 Halomonas
29	50	2.0	1134	1	AB014924	AB014924 Halomonas
C	49.6	2.0	339650	1	SCO939108	AL939108 Streptomy
31	49	2.0	749	14	NPORPT2	DI3755 Orygia psuu
C	49	2.0	14483	1	AE006992	AE006992 Mycobacte
33	49	2.0	67200	1	MTV017	AL021897 Mycobacte
34	49	2.0	131995	14	OPU75930	U75930 Orygia psuu
C	49	2.0	314100	1	SCO939106	AL939106 Streptomy
36	48.6	2.0	41782	1	AY118081	AY118081 Streptomy
C	48.6	2.0	277000	1	SCO939109	AL939109 Streptomy
38	48.4	2.0	16150	1	MTV044	AL021989 Mycobacte
39	48.4	2.0	16168	1	AE006985	AE006985 Mycobacte
40	48.4	2.0	299925	1	AP005048	AP005048 Streptomy
41	48.2	1.9	141942	2	AC130725	AC130725 Oryza sat
C	48.2	1.9	158911	2	AC136492	AC136492 Oryza sat
43	48	1.9	300100	1	SCO939123	AL939123 Streptomy
44	48	1.9	10522	1	AE004514	AE004514 Pseudomon
C	47.8	1.9	10522	1	AE004514	AE004514 Pseudomon

ALIGNMENTS

RESULT 1
ARI25096
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

ARI25096
Sequence 47 from patent US 6177075.
ARI25096
GI:14111158
Unknown.
Unclassified.
1 (bases 1 to 2478)
Christian,P.Daniel.,
Hanzlik,T.Nelson.
Insect viruses and their uses in protecting plants
Patent: US 6177075-A 47 23-JAN-2001;

2478 bp
DNA
linear
PAT 16-MAY-2001

FEATURES		Location/Qualifiers	
source	1. .2478		
BASE COUNT	516 a 854 c 607 g 501 t		
ORIGIN			
Query Match 100.0%; Score 2478; DB 6; Length 2478;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GTCTTTCTTTCTTTACCAAGTGTGTAATAATTTAAACAAAGAGAAAACCCAGGACCGTAA	60
Db	1	GTCTTTCTTTCTTTACCAAGTGTGTAATAATTTAAACAAAGAGAAAACCCAGGACCGTAA	60
Qy	61	CCCGCCCTTACACACCTCGAGTCGAGTCGACACCGGATTATACGTCGCGCCACACACGGC	120
Db	61	CCCGCCCTTACACACCTCGAGTCGAGTCGACACCGGATTATACGTCGCGCCACACACGGC	120
Qy	121	GCCTTTTCCGACCACTCTCGAGAGTCGTTGGGAGTTTCGTCGTCGACCAACCCGGTTGGCA	180
Db	121	GCCTTTTCCGACCACTCTCGAGAGTCGTTGGGAGTTTCGTCGTCGACCAACCCGGTTGGCA	180
Qy	181	GTGACAGAGCGTTCCGGACCACTAGAACCTCTCGAGCGACGACACACAGCACACA	240
Db	181	GTGACAGAGCGTTCCGGACCACTAGAACCTCTCGAGCGACGACACACAGCACACA	240
Qy	241	CGCCCTTAGCTGACCTACCGCAGCGTTGATAGCGCGGATTATGAGCGAGCACACCATC	300
Db	241	CGCCCTTAGCTGACCTACCGCAGCGTTGATAGCGCGGATTATGAGCGAGCACACCATC	300
Qy	301	GCCCACTCCATCACATTACCAACCCGGTTACACCTTGCCTTAATACCCCTGAACTGAA	360
Db	301	GCCCACTCCATCACATTACCAACCCGGTTACACCTTGCCTTAATACCCCTGAACTGAA	360
Qy	361	CGAGATGGGAGATGCTGGAGTGGCGTCAAGCGACTCAACACCGTCGCGGAAACCCGTA	420
Db	361	CGAGATGGGAGATGCTGGAGTGGCGTCAAGCGACTCAACACCGTCGCGGAAACCCGTA	420
Qy	421	ACGTTCCGGTCAGCGCCACACCGTCACCGTCAATGGTAGAGAAACCAACCGGCTCGGA	480
Db	421	ACGTTCCGGTCAGCGCCACACCGTCACCGTCAATGGTAGAGAAACCAACCGGCTCGGA	480
Qy	481	CCGGAAGGCAAGTTTCTCCCTCGACAAATTTCCACCGCTGCTGCAACAGACCTCGCGCAA	540
Db	481	CCGGAAGGCAAGTTTCTCCCTCGACAAATTTCCACCGCTGCTGCAACAGACCTCGCGCAA	540
Qy	541	GCCTTGACGCCAACAACCGTCACTTTCCCGCTAAACATCTAGATGATGCCGAAATTCGGA	600
Db	541	GCCTTGACGCCAACAACCGTCACTTTCCCGCTAAACATCTAGATGATGCCGAAATTCGGA	600
Qy	601	ATTGGGCCAAGGGAAGATCGACCTCGACTCGGATTCGATCGGCTGCTTCAAGTACC	660
Db	601	ATTGGGCCAAGGGAAGATCGACCTCGACTCGGATTCGATCGGCTGCTTCAAGTACC	660
Qy	661	TTGACCCAGCGGCTGACAGAGTCTGGCGCGCGCTCGGCGAGTACTCGAAGATCCCTG	720
Db	661	TTGACCCAGCGGCTGACAGAGTCTGGCGCGCGCTCGGCGAGTACTCGAAGATCCCTG	720
Qy	721	ACGGCTCGTCAAGTCTCCGTCGACGAGATGAAGAGATCTATTAACCGAGGAGTGCC	780
Db	721	ACGGCTCGTCAAGTCTCCGTCGACGAGATGAAGAGATCTATTAACCGAGGAGTGCC	780
Qy	781	CCGTGCTCACTGACGTGCTCCGTCGACCGCGCGCGCTCGAGTCTCGATTTTCT	840
Db	781	CCGTGCTCACTGACGTGCTCCGTCGACCGCGCGCGCTCGAGTCTCGATTTTCT	840
Qy	841	CTTTTCCGATGTTTCAAGACCGCTACGTCGCGCTAGCGAACCGTCAAGAGATGT	900
Db	841	CTTTTCCGATGTTTCAAGACCGCTACGTCGCGCTAGCGAACCGTCAAGAGATGT	900
Qy	901	CGCTCGAGTGTGTCACGACCTCATCGATGGCTCAACAACTCTCGCGAGCTGGCGTTATG	960
Db	901	CGCTCGAGTGTGTCACGACCTCATCGATGGCTCAACAACTCTCGCGAGCTGGCGTTATG	960

Qy	961	TCGTTGACTCTGAACAGTGGATTAACTTCAACAAATGACACACCGTACTACGTCGCAATCC	1020
Db	961	TCGTTGACTCTGAACAGTGGATTAACTTCAACAAATGACACACCGTACTACGTCGCAATCC	1020
Qy	1021	GCCTTCTAGTCCAAACCTACGAGTTCGAGACCCACAGAGGGCTTGTTCGCAAGTCT	1080
Db	1021	GCCTTCTAGTCCAAACCTACGAGTTCGAGACCCACAGAGGGCTTGTTCGCAAGTCT	1080
Qy	1081	CAGACTACCGCTCACATTATAAGGCGATACATGTGAAGCCAAATGCAACACTCTGTCG	1140
Db	1081	CAGACTACCGCTCACATTATAAGGCGATACATGTGAAGCCAAATGCAACACTCTGTCG	1140
Qy	1141	ACCAAGGCTTTTGGATCGGCGCCAGTACGCTCTCACCCCGACTAGCTACCGCAGTACG	1200
Db	1141	ACCAAGGCTTTTGGATCGGCGCCAGTACGCTCTCACCCCGACTAGCTACCGCAGTACG	1200
Qy	1201	ACGTACGAGGAGCTACGCTCTGCACACTTTTGACCTTCGCCAGACCAATCCAGCGCGGTG	1260
Db	1201	ACGTACGAGGAGCTACGCTCTGCACACTTTTGACCTTCGCCAGACCAATCCAGCGCGGTG	1260
Qy	1261	CACCTCGGTTTGTGTCGGCAGGTTTGCACAGGGTGGCAGTTCGCGCTGACGCACTCCAG	1320
Db	1261	CACCTCGGTTTGTGTCGGCAGGTTTGCACAGGGTGGCAGTTCGCGCTGACGCACTCCAG	1320
Qy	1321	CCTGGGAGCAGGCACTCCTCGGCTGCTACCTCACCTGGCGCCACAACCGGTACTACTTTCC	1380
Db	1321	CCTGGGAGCAGGCACTCCTCGGCTGCTACCTCACCTGGCGCCACAACCGGTACTACTTTCC	1380
Qy	1381	CAGTGGCTCGTGTAGTACGTTCTCCTGAGGTTTTCGCCCTTGGAGCGCTACGACCCGA	1440
Db	1381	CAGTGGCTCGTGTAGTACGTTCTCCTGAGGTTTTCGCCCTTGGAGCGCTACGACCCGA	1440
Qy	1441	ACGAGCGCTCTTGACACGACTTTCGCTTCGCGAGAGACACCGTCACTTTCGGGAGGTCG	1500
Db	1441	ACGAGCGCTCTTGACACGACTTTCGCTTCGCGAGAGACACCGTCACTTTCGGGAGGTCG	1500
Qy	1501	CCGTGCGAGAGTCTGTGTGACCAACACCCCGCGCGGCGGAGCGCCGCCACCTTCA	1560
Db	1501	CCGTGCGAGAGTCTGTGTGACCAACACCCCGCGCGGCGGAGCGCCGCCACCTTCA	1560
Qy	1561	CCGTGAGAGTCCCGCTTCAACACGCTTACCAACACGCTGTTTGAACACGCTCTTAG	1620
Db	1561	CCGTGAGAGTCCCGCTTCAACACGCTTACCAACACGCTGTTTGAACACGCTCTTAG	1620
Qy	1621	AGACTCGACCTCTCTCGTAGGCTCGAACTCCCTATGCCACCTGCTGACCTTGGACAGA	1680
Db	1621	AGACTCGACCTCTCTCGTAGGCTCGAACTCCCTATGCCACCTGCTGACCTTGGACAGA	1680
Qy	1681	CCGTGCGCCAAACACCCGAAAGATCGAGCAGTCTTCTTAAAGAAACACTTGGCTGCTATT	1740
Db	1681	CCGTGCGCCAAACACCCGAAAGATCGAGCAGTCTTCTTAAAGAAACACTTGGCTGCTATT	1740
Qy	1741	TGGTCCACTCCAAATGCGAAACCCGTTTTCAGCTCAACGCCAGCAGCTCTTTGGCG	1800
Db	1741	TGGTCCACTCCAAATGCGAAACCCGTTTTCAGCTCAACGCCAGCAGCTCTTTGGCG	1800
Qy	1801	CCGTTCCTTCAACAAATCCGGGTTATGAGCGCACACCGGACCTCCCGGACTACACTGGCA	1860
Db	1801	CCGTTCCTTCAACAAATCCGGGTTATGAGCGCACACCGGACCTCCCGGACTACACTGGCA	1860
Qy	1861	TCCGTGACTCAATCGACCAAGAAATGTCCACCGCTGTGGCCCACTTCCGCTCACTCTCCC	1920
Db	1861	TCCGTGACTCAATCGACCAAGAAATGTCCACCGCTGTGGCCCACTTCCGCTCACTCTCCC	1920
Qy	1921	ACTCCTGAGTATCGTCACTAAGACCTACAGGGTGGGAAGGGCTCAAGAACGTCACAA	1980
Db	1921	ACTCCTGAGTATCGTCACTAAGACCTACAGGGTGGGAAGGGCTCAAGAACGTCACAA	1980
Qy	1981	CGCTTTCCGCAATTCGCGCACCGGGCTCTCCAGAAATGAGGAGATCTCTGCGCTCG	2040
Db	1981	CGCTTTCCGCAATTCGCGCACCGGGCTCTCTCCAGAAATGAGGAGATCTCTGCGCTCG	2040

1321 CTTGGAGAGGAGCATCTCGGTGGCTACCTCACCTGGCGGCACAAACGGTACTACTTTCC 1380
1321 CTTGGAGAGGAGCATCTCGGTGGCTACCTCACCTGGCGGCACAAACGGTACTACTTTCC 1380
1381 CAGCTGGCTCGGTAGTACGTTCTCTGAGGGTTTCGCCCTTGAGCGCTACGACCCGA 1440
1381 CAGCTGGCTCGGTAGTACGTTCTCTGAGGGTTTCGCCCTTGAGCGCTACGACCCGA 1440
1441 ACGAGGCTCTTGGACCGACATCTCGCTTCGGAGGAGACACCGTCACTTTCCGGGAGTGC 1500
1441 ACGAGGCTCTTGGACCGACATCTCGCTTCGGAGGAGACACCGTCACTTTCCGGGAGTGC 1500
1501 CCGTCGACGAGTCTGTGTGACCAACACCCCGCGGGGCGGACGCCGCCACCTTCA 1560
1501 CCGTCGACGAGTCTGTGTGACCAACACCCCGCGGGGCGGACGCCGCCACCTTCA 1560
1561 CCGTGAGAGTGCCTCTTCAACCGCTTACCAACACCGCTGTTTAGGAACACGCTCTTAG 1620
1561 CCGTGAGAGTGCCTCTTCAACCGCTTACCAACACCGCTGTTTAGGAACACGCTCTTAG 1620
1621 AGACTCGACCTCTCTCTGAGGCTCGAATCCCTATGCCACCTGCTGACTTTGGACAGA 1680
1621 AGACTCGACCTCTCTCTGAGGCTCGAATCCCTATGCCACCTGCTGACTTTGGACAGA 1680
1681 CCGTCGCCAACAACCCGGAAGATCGAGCAGTCTTCTTAAAGAAACACTTGGCTGCTATT 1740
1681 CCGTCGCCAACAACCCGGAAGATCGAGCAGTCTTCTTAAAGAAACACTTGGCTGCTATT 1740
1741 TGGTCCACTCCAAATATGGAAACCCGCTTTTCCAGCTCACGCCAGCCAGCTCTTTGGCG 1800
1741 TGGTCCACTCCAAATATGGAAACCCGCTTTTCCAGCTCACGCCAGCCAGCTCTTTGGCG 1800
1801 CCGTTTCTTCAACATCGGGTTATGAGCGCACACGCGACCTCCCGACTACACTGGCA 1860
1801 CCGTTTCTTCAACATCGGGTTATGAGCGCACACGCGACCTCCCGACTACACTGGCA 1860
1861 TCCGTGACTCATTTCCAGCAGAAATGTCACCGCTGTGGCCACATTCGCTCACTCTCCC 1920
1861 TCCGTGACTCATTTCCAGCAGAAATGTCACCGCTGTGGCCACATTCGCTCACTCTCCC 1920
1921 ACTCTGAGTATCGTCACTAAGACCTTACAGGGTTGGGAAGGCGTCAAGACGTCACAA 1980
1921 ACTCTGAGTATCGTCACTAAGACCTTACAGGGTTGGGAAGGCGTCAAGACGTCACAA 1980
1981 CGCTTTCCGCCAATTCGCGACGCGGGCTCTCAAGAAAGAGAGATCTCTGCTGCG 2040
1981 CGCTTTCCGCCAATTCGCGACGCGGGCTCTCTCAAGAAAGAGAGATCTCTGCTGCG 2040
2041 CCGAGACTGSCCACCCTCTCAAGTGTCTACCCGCGACTGACAACTTCCGGGCGG 2100
2041 CCGAGACTGSCCACCCTCTCAAGTGTCTACCCGCGACTGACAACTTCCGGGCGG 2100
2101 CCGTTTCTGCTTCCGCGGAACATGCTCTCTCGTGTGCTCGTGTGAAGTCGAGGCAACGTCT 2160
2101 CCGTTTCTGCTTCCGCGGAACATGCTCTCTCGTGTGCTCGTGTGAAGTCGAGGCAACGTCT 2160
2161 CCATCATCAAGTCTGCTGAGAGATGCGCTCGGCGGGCTCAGTCCGGCTCGGAGC 2220
2161 CCATCATCAAGTCTGCTGAGAGATGCGCTCGGCGGGCTCAGTCCGGCTCGGAGC 2220
2221 TACCCGAGTCTGATAGTGTACCAAGGAGATTGCGCGCTGTCGCGCGCGCGAG 2280
2221 TACCCGAGTCTGATAGTGTACCAAGGAGATTGCGCGCTGTCGCGCGCGCGAG 2280
2281 CGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
2281 CGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
2341 ACGGCGTGGTCCCGACATTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
2341 ACGGCGTGGTCCCGACATTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
2401 CGAGGGTCTGCGCGGGTTTCGATTCCATTCCCAAGCGGCAAGAGGACGTAGTTAGCTCT 2460

2401 CGAGGGTCTGCGCGGGTTTCGATTCCATTCCCAAGCGGCAAGAGGACGTAGTTAGCTCT 2460
2461 GCCTCCCTCGGGATACCA 2478
2461 GCCTCCCTCGGGATACCA 2478
RESULT 3
AR159968
LOCUS AR159968 2478 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6251654.
ACCESSION AR159968
VERSION AR159968.1 GI:16222852
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2478)
AUTHORS Gordon, K. Heinrich, and Hanzlik, T. Nelson.
TITLE Modified small RNA viruses
JOURNAL Patent: US 6251654-A 1 26-JUN-2001;
FEATURES Location/Qualifiers
source 1..2478
BASE COUNT 516 a 854 c 607 g 501 t
ORIGIN
Query Match 100.0%; Score 2478; DB 6; Length 2478;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTTTTCTTTTACCAAGTGTGTAATAAACAAGAAAGAAACACGACCGTAA 60
DB 1 GTTTTCTTTTACCAAGTGTGTAATAAACAAGAAAGAAACACGACCGTAA 60
QY 61 CCCGGCCTTACACACCTCGAGTCCGTGACACCGGATTATACGTCGCCACACACGCG 120
DB 61 CCCGGCCTTACACACCTCGAGTCCGTGACACCGGATTATACGTCGCCACACACGCG 120
QY 121 GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTTCTCGTGACACCCGGTTGGCA 180
DB 121 GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTTCTCGTGACACCCGGTTGGCA 180
QY 181 GTCCACAGAGCGTTCCGACCCACTAGAACCTCTCCGAGCGACGACACACAGCACACA 240
DB 181 GTCCACAGAGCGTTCCGACCCACTAGAACCTCTCCGAGCGACGACACACAGCACACA 240
QY 241 CCGCCTTAGCTGCACTACCGCAGCTTGTATAGCGCGGATTTATGAGCGAGCACACCATC 300
DB 241 CCGCCTTAGCTGCACTACCGCAGCTTGTATAGCGCGGATTTATGAGCGAGCACACCATC 300
QY 301 GCCCACTCCATTCACATTCACACCGGGTTACACCTTGCCTTAATACCCCTGAACTCGAA 360
DB 301 GCCCACTCCATTCACATTCACACCGGGTTACACCTTGCCTTAATACCCCTGAACTCGAA 360
QY 361 GCAGGATGGGAGATGCTGGAGTGGCGTCAAGCCCTCAACCCGTCGCGGAAACCCGTA 420
DB 361 GCAGGATGGGAGATGCTGGAGTGGCGTCAAGCCCTCAACCCGTCGCGGAAACCCGTA 420
QY 421 ACCTTCCGGTTCAGCGCCCAACACCGCTCACCGCTCAATGGTAGAAGAAACCAACCGGTCGGA 480
DB 421 ACCTTCCGGTTCAGCGCCCAACACCGCTCACCGCTCAATGGTAGAAGAAACCAACCGGTCGGA 480
QY 481 CCGGAAAGCAAGTTTCTCCCTGCAATTTCAACCGTGTGCAAGACCTCGGCGAAA 540
DB 481 CCGGAAAGCAAGTTTCTCCCTGCAATTTCAACCGTGTGCAAGACCTCGGCGAAA 540
QY 541 GCCTTGGAGCGCAACACCGTCACTTTCCCGCTACATCTCTAGCATGCCGAAATTCGGA 600
DB 541 GCCTTGGAGCGCAACACCGTCACTTTCCCGCTACATCTCTAGCATGCCGAAATTCGGA 600
QY 601 ATTGGGCCAAGGAAAGATCGACCTCGACTCCGATTCATCCATCGGCTGGTACTTCAAGTACC 660


```
/protein_id="AAC37885.1"
/db_xref="GI:609505"
/translation="MGDAGVASQRPNNRRGTRNVRVSANTVTVNGRRNRORRTGRQVS
PPDNFTAAQDLAOLSDANTVTFPANISSMPEFRNWKIDLDSDLSIGWFKYLDPA
GATESARAVGSKIPDGLVKFSVDARELYNEECPVVTDVSVPLDGRGWSLIFSP
PMFTAVAVANVENKMSLDVNDLIEWANLADRWVVDSEQMINFTNDVTYVRI
RVLRTYDVPDTEGLVTRSDYRLTYKAITCEANMPLVDQGGWIGQYALYTLSP
QYDVEAVALHTLTPARSSAAALAFVWAGLPQGGTAPAGTAPMEQSSGGYLTPRHN
GTTTPAGSVYLVPEGFALERYDNDGSMWTFASAGDTVTTPRVADEVVVNNPAGG
GSATPTFVRVPSNAYTNTVFRNTLLETSPRSRLLELPMFPADFCQTVANNPKIEQSL
LKEHLGCVLVHSKRNVPQVLTTPASSFGAVSNPNPGBYERDLDPDYTGIRDSFQDNMS
TAVAHFSLSHSCSLVTKYQGWGVNTVNTPCOPAHAGLLKNEELICLADDLATRL
TGVDPATDNFAAASAFANMLSSVLKSEATSSIKSVGTAVGAHQSLAKLPLGLLM
SVPGKIAARVARRARRRRAAN"
1497..1547
/stem_loop
/gene="p71"
/note="this region is found deleted in cDNA product after
reverse transcription by MuMolV RT or its derivative,
Superscript I."
/evidence=experimental
2310..2478
/feature="putative"
/feature="putative"
/misc_structure 2326..2478
/standard_name="viral tRNA-like structure"
/feature="tRNA-like structure without pseudo knot can be made
from this region, has valine anticodon, with 30 bases in
variable loop region; putative"
/function="unknown"
BASE COUNT 516 a 854 c 607 g 501 t
ORIGIN
Query Match 100.0%; Score 2478; DB 14; Length 2478;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTTTTCTTTTACCAAGTGTGTAAATTTAAACAAGAGAAAACCGAGCCGTAA 60
DB 1 GTTTTCTTTTACCAAGTGTGTAAATTTAAACAAGAGAAAACCGAGCCGTAA 60
QY 61 CCCGCCCTTACACACCTCGAGTCGTCACACCGGATTATACGTCGCCACCAACCGC 120
DB 61 CCCGCCCTTACACACCTCGAGTCGTCACACCGGATTATACGTCGCCACCAACCGC 120
QY 121 GCCTTTTCCGACCACTCTCGAGAGTCGTTGGAGTTTCGTCGTCACCAACCGGTGGCA 180
DB 121 GCCTTTTCCGACCACTCTCGAGAGTCGTTGGAGTTTCGTCGTCACCAACCGGTGGCA 180
QY 181 GTCGACAGACGCTTCGGACCACTAGAACCTCTCTCGAGGAGCGACACAGCACACA 240
DB 181 GTCGACAGACGCTTCGGACCACTAGAACCTCTCTCGAGGAGCGACACAGCACACA 240
QY 241 CGGCTTAGCTGCACCTAGCGAGCGTTGATAGCGCGGATTATGAGGAGGACACCATC 300
DB 241 CGGCTTAGCTGCACCTAGCGAGCGTTGATAGCGCGGATTATGAGGAGGACACCATC 300
QY 301 GCCCACTCCATCACATTACACCCGGTTACACCTTGGCCCTTAATACCCCTGAACTGAA 360
DB 301 GCCCACTCCATCACATTACACCCGGTTACACCTTGGCCCTTAATACCCCTGAACTGAA 360
QY 361 CGAGATGGGAGATGCTGAGAGTGGCGTCAAGCGACCTTCAACCGTCGCGGAACCCGTA 420
DB 361 CGAGATGGGAGATGCTGAGAGTGGCGTCAAGCGACCTTCAACCGTCGCGGAACCCGTA 420
QY 421 AGTTTCGGGTGAGCGGACCAACCGTCACCGTCAATGGTAGAGAAACCAACGGCGTCGA 480
DB 421 AGTTTCGGGTGAGCGGACCAACCGTCACCGTCAATGGTAGAGAAACCAACGGCGTCGA 480
QY 481 CCGGAAGGCAAGTTTCTCCCTCTGACAAATTTTACCGCTGCTGCACAAAGACCTCGCGCAA 540
DB 481 CCGGAAGGCAAGTTTCTCCCTCTGACAAATTTTACCGCTGCTGCACAAAGACCTCGCGCAA 540
QY 541 GCCTTGACGCCAACCGGTCACTTTCCCGCTTAACATCTCTAGCATGCCCGAATTCGGA 600
DB 541 GCCTTGACGCCAACCGGTCACTTTCCCGCTTAACATCTCTAGCATGCCCGAATTCGGA 600
```

```
Db 541 GCCTTGACGCCAACACCGGTCACTTTCCCGCTAAATCTCTAGCATGCCCGAATTCGGA 600
QY 601 ATTGGGCAAGGGAAAGATCGACCTCGACTCCGATTCCATCGGCTGGTACTTCAAGTACC 660
DB 601 ATTGGGCAAGGGAAAGATCGACCTCGACTCCGATTCCATCGGCTGGTACTTCAAGTACC 660
QY 661 TTGACCCAGCGGGTGCTACAGAGTCTGGCGCGCGCTCGGCGAGTACTTCGAAGATCCCTG 720
DB 661 TTGACCCAGCGGGTGCTACAGAGTCTGGCGCGCGCTCGGCGAGTACTTCGAAGATCCCTG 720
QY 721 ACAGCTCGTCAAGTCTTCGCTCGACGAGAGATAAGAGAGATCTATACGAGGAGTACC 780
DB 721 ACAGCTCGTCAAGTCTTCGCTCGACGAGAGATAAGAGAGATCTATACGAGGAGTACC 780
QY 781 CCGTCTGCTACGTGCTCGTCTCCCTCGACGCGCGCTCGAGTGGAGCTCTTCGATTTTCT 840
DB 781 CCGTCTGCTACGTGCTCGTCTCCCTCGACGCGCGCTCGAGTGGAGCTCTTCGATTTTCT 840
QY 841 CTTTCCGATGTTTCAAGAACCGCTTACGTGCGCGCTAGCGAACGTTCGAGAACGAGATGT 900
DB 841 CTTTCCGATGTTTCAAGAACCGCTTACGTGCGCGCTAGCGAACGTTCGAGAACGAGATGT 900
QY 901 CGCTCGAGCTTGTCAACGAGCTCATCGAGTGGCTCAACAATCTCGCCGACTGGCGTTATG 960
DB 901 CGCTCGAGCTTGTCAACGAGCTCATCGAGTGGCTCAACAATCTCGCCGACTGGCGTTATG 960
QY 961 TCGTTGACTCTGAACAGTGGATTAACTTCAACAATGACACCACTACTACGTCGCGATCC 1020
DB 961 TCGTTGACTCTGAACAGTGGATTAACTTCAACAATGACACCACTACTACGTCGCGATCC 1020
QY 1021 CGGTTCTAGTCCAAACCTTACGAGTTCAGACCCCAACGAGGCGCTTGTTCGACAGTCT 1080
DB 1021 CGGTTCTAGTCCAAACCTTACGAGTTCAGACCCCAACGAGGCGCTTGTTCGACAGTCT 1080
QY 1081 CAGACTACCGCTCACATTTAAGGCGATAACATGTGAAGCCAAATGTCGAGTCTGCTG 1140
DB 1081 CAGACTACCGCTCACATTTAAGGCGATAACATGTGAAGCCAAATGTCGAGTCTGCTG 1140
QY 1141 ACCAAGGCTTTGGATCGGCGCCAGTACGCTCTCACCCCGACTAGCTACCGCAGTACG 1200
DB 1141 ACCAAGGCTTTGGATCGGCGCCAGTACGCTCTCACCCCGACTAGCTACCGCAGTACG 1200
QY 1201 ACGTGACGAGGCTTACGCTCTGACACTTTTGA CTTTGGCGCAGACCATTCAGCGCGCTG 1260
DB 1201 ACGTGACGAGGCTTACGCTCTGACACTTTTGA CTTTGGCGCAGACCATTCAGCGCGCTG 1260
QY 1261 CACTCGGCTTTGTGGGCGAGTTTGGCAGAGGTTGGCAGGCTGCGCTGCGAGGCTCCAG 1320
DB 1261 CACTCGGCTTTGTGGGCGAGTTTGGCAGAGGTTGGCAGGCTGCGCTGCGAGGCTCCAG 1320
QY 1321 CCTGGAGCAGGCTATCTCGGCTGGCTACCTCACCTGGCGCCAAACGGTACTACTTTCC 1380
DB 1321 CCTGGAGCAGGCTATCTCGGCTGGCTACCTCACCTGGCGCCAAACGGTACTACTTTCC 1380
QY 1381 CAGTGGCTCCGTTAGCTACGTTCTCGCTGAGGTTTGGCGCTTGGAGCGCTACGACCGGA 1440
DB 1381 CAGTGGCTCCGTTAGCTACGTTCTCGCTGAGGTTTGGCGCTTGGAGCGCTACGACCGGA 1440
QY 1441 ACAGCGCTCTTGAACCGACTTTCGCTTCGCGAGGAGACACCGTCACTTTCCGCGAGGTCG 1500
DB 1441 ACAGCGCTCTTGAACCGACTTTCGCTTCGCGAGGAGACACCGTCACTTTCCGCGAGGTCG 1500
QY 1501 CCGTTCGAGGCTCGTTGTGACCAACACCCCGCGCGGCGGAGCGCCGCCACTTCA 1560
DB 1501 CCGTTCGAGGCTCGTTGTGACCAACACCCCGCGCGGCGGAGCGCCGCCACTTCA 1560
QY 1561 CCGTTCGAGGCTCGCCCTTTCAAACGCTTACACCAACACCGCTTTTAGGAACACGCTCTTAG 1620
DB 1561 CCGTTCGAGGCTCGCCCTTTCAAACGCTTACACCAACACCGCTTTTAGGAACACGCTCTTAG 1620
QY 1621 AGACTCGACCTCTCTGCTAGGCTCGAACTCCCTATGCGACCTGCTGACTTTGACAGAGA 1680
DB 1621 AGACTCGACCTCTCTGCTAGGCTCGAACTCCCTATGCGACCTGCTGACTTTGACAGAGA 1680
```

QY 1681 CGGTGCGCAACAAACCGAAGATCGAGAGTGGCTTCTTTAAAGAAACACTTGGCTGCTATT 1740
DB 1681 CGGTGCGCAACAAACCGAAGATCGAGAGTGGCTTCTTTAAAGAAACACTTGGCTGCTATT 1740
QY 1741 TGGTCCACTCCAAATGCGAAACCCGGTTTCCAGCTCACGCCAGCCAGCTCTTTGGCG 1800
DB 1741 TGGTCCACTCCAAATGCGAAACCCGGTTTCCAGCTCACGCCAGCCAGCTCTTTGGCG 1800
QY 1801 CGGTTTCTTCAACAAATCGGGTTATGAGCGCACACGCGACCTCCCGACTACACTGGCA 1860
DB 1801 CGGTTTCTTCAACAAATCGGGTTATGAGCGCACACGCGACCTCCCGACTACACTGGCA 1860
QY 1861 TCCGTGACTCATTCGACAGAACATGTCACCGCTGTGCGCCACTTTCGGTCACTCTCCC 1920
DB 1861 TCCGTGACTCATTCGACAGAACATGTCACCGCTGTGCGCCACTTTCGGTCACTCTCCC 1920
QY 1921 ACTCTCGAGTATCGTCACTAAGACTTACAGGGTTGGGAAGGCGTACGAACGTCACAA 1980
DB 1921 ACTCTCGAGTATCGTCACTAAGACTTACAGGGTTGGGAAGGCGTACGAACGTCACAA 1980
QY 1981 CGCTTTTCGGGCAATTCGGGCAACGGCGGCTCTCAAGATGAGAGATCTCTTCGCTCG 2040
DB 1981 CGCTTTTCGGGCAATTCGGGCAACGGCGGCTCTCTCAAGATGAGAGATCTCTTCGCTCG 2040
QY 2041 CCGAGACTTGGCCACCCGCTCTCACAGGTGTCTACCCCGCCACTGACAACTTCGGGCGG 2100
DB 2041 CCGAGACTTGGCCACCCGCTCTCACAGGTGTCTACCCCGCCACTGACAACTTCGGGCGG 2100
QY 2101 CCGTTTCTGCTTCCGCCGGAACATGCTGTCTCCGCTGCTGAAGTCGGAGGCAAGCTCT 2160
DB 2101 CCGTTTCTGCTTCCGCCGGAACATGCTGTCTCCGCTGCTGAAGTCGGAGGCAAGCTCT 2160
QY 2161 CCATCATCAAGTCCGTTGGGAGACTCGCTGCGGCGGCTCAGTCCGGCTTCGGGAGC 2220
DB 2161 CCATCATCAAGTCCGTTGGGAGACTCGCTGCGGCGGCTCAGTCCGGCTTCGGGAGC 2220
QY 2221 TACCGGACTGCTAATAGTGTACCAAGGAAGATTGCCGCGCTGTCGCGCGCGCGAG 2280
DB 2221 TACCGGACTGCTAATAGTGTACCAAGGAAGATTGCCGCGCTGTCGCGCGCGCGAG 2280
QY 2281 CGCGCGCGCGCGCTCGTGCATTAAGTTTGTCTGCTCTCTGTTTCCGCTTTGCTAAA 2340
DB 2281 CGCGCGCGCGCGCTCGTGCATTAAGTTTGTCTGCTCTCTGTTTCCGCTTTGCTAAA 2340
QY 2341 ACGGCGTGGTCCCGCACATTAAGCTACCTAAAGACTCTGAGTCCCGTCTGTTACA 2400
DB 2341 ACGGCGTGGTCCCGCACATTAAGCTACCTAAAGACTCTGAGTCCCGTCTGTTACA 2400
QY 2401 CGAGGGTCTCGCGGGTTCCGATTCATTCGCAAGCGGCAAGAGAGTAGTGTCT 2460
DB 2401 CGAGGGTCTCGCGGGTTCCGATTCATTCGCAAGCGGCAAGAGAGTAGTGTCT 2460
QY 2461 GCGTCCCTCGGATACCA 2478
DB 2461 GCGTCCCTCGGATACCA 2478

RESULT 7
LOCUS AR125098 2479 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 51 from patent US 6177075.
ACCESSION AR125098
VERSION AR125098.1 GI:14111160
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2479)
AUTHORS Christian,P.Daniel., Gordon,K.Hienrich.Julius. and Hanzlik,T.Nelson.
TITLE Insect viruses and their uses in protecting plants
JOURNAL Patent: US 6177075-A 51 23-JAN-2001;

FEATURES
Source

Location/Qualifiers
1. .2479
/organism="unknown"

BASE COUNT 516 a 855 c 607 g 501 t
ORIGIN

Query Match 99.6%; Score 2467; DB 6; Length 2479;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2478; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GTTTTCTTTCTTTACCAAGTGTGTAATAAATTTAAACAAAGAAAGAAACCGAGACCGTAA 60
DB 1 GTTTTCTTTCTTTACCAAGTGTGTAATAAATTTAAACAAAGAAAGAAACCGAGACCGTAA 60
QY 61 CCGGCGCTTTACACACCTCGAGTCCGTGACCAACCGGATTTATACGTCCGCCACACACACGGC 120
DB 61 CCGGCGCTTTACACACCTCGAGTCCGTGACCAACCGGATTTATACGTCCGCCACACACGGC 120
QY 121 GCCTTTTCCGACCACTCTCGAGAGTCTGTGGAGTTTCTGTCGCTGACACACCGGTTGGCA 180
DB 121 GCCTTTTCCGACCACTCTCGAGAGTCTGTGGAGTTTCTGTCGCTGACACACCGGTTGGCA 180
QY 181 GTGACAGACGCTTCCGGACCACTAGAACTCTTCGAGGACGACACACACAGCACACACA 240
DB 181 GTGACAGACGCTTCCGGACCACTAGAACTCTTCGAGGACGACACACACAGCACACACA 240
QY 241 CCGCTTGTAGTGTACCTTACCGGAGCTTGTAGCGCGGATTTATAGCGGAGCACACCATC 300
DB 241 CCGCTTGTAGTGTACCTTACCGGAGCTTGTAGCGCGGATTTATAGCGGAGCACACCATC 300
QY 301 GCCCACTCCATTCACATTTACCAACCGGTTTACACCTTGCCTTAATACCCCTTGAACCTGAA 360
DB 301 GCCCACTCCATTCACATTTACCAACCGGTTTACACCTTGCCTTAATACCCCTTGAACCTGAA 360
QY 361 GCAGGATGGAGATGCTTGGAGTGGCTTACAGGACCTTCAACCGCTCGCGAACCCTGA 420
DB 361 GCAGGATGGAGATGCTTGGAGTGGCTTACAGGACCTTCAACCGCTCGCGAACCCTGA 420
QY 421 ACCTTCCGGTTCAGCGCAACACCGTTCACCGTCAATGTAGAGAAACCAACCGGCTCGGA 480
DB 421 ACCTTCCGGTTCAGCGCAACACCGTTCACCGTCAATGTAGAGAAACCAACCGGCTCGGA 480
QY 481 CCGGAAGCGCAAGTTTCTCCCTTGCACAAATTTACCGCTGTGTGCAACAGACTCTCGCAAA 540
DB 481 CCGGAAGCGCAAGTTTCTCCCTTGCACAAATTTACCGCTGTGTGCAACAGACTCTCGCAAA 540
QY 541 GCCTTGAACGCAACACCGTCTACTTT - CCGCGTAAACATCTCTAGCATGCCGAATTCGG 599
DB 541 GCCTTGAACGCAACACCGTCTACTTTCCCGCTTAAACATCTCTAGCATGCCGAATTCGG 600
QY 600 AATTGGGCGCAAGGAAGATCGACCTCGACTCCGATTCATCCGCTGCTACTTCAAGTAC 659
DB 600 AATTGGGCGCAAGGAAGATCGACCTCGACTCCGATTCATCCGCTGCTACTTCAAGTAC 659
QY 660 CTTGACCCAGCGGTTCTACAGAGTCTGCGCGCGCTGCGCGAGTACTCGAAGATCCCT 719
DB 660 CTTGACCCAGCGGTTCTACAGAGTCTGCGCGCGCTGCGCGAGTACTCGAAGATCCCT 720
QY 720 GACGCGCTCGTCAAGTCTTCTCGTGCAGGAGATTAAGAGAGATCTATAACGAGAGTGC 779
DB 720 GACGCGCTCGTCAAGTCTTCTCGTGCAGGAGATTAAGAGAGATCTATAACGAGAGTGC 780
QY 780 CCGCTGCTCACTGAGTGTCCCTCCCTCGAGCGCGCGAGTGGAGCTCTCGATTTTC 839
DB 780 CCGCTGCTCACTGAGTGTCCCTCCCTCGAGCGCGCGAGTGGAGCTCTCGATTTTC 840
QY 840 TCCCTTTCCGATGTTTCAAGAACCGCTTACGTCGCGTACGTCGAGAACCAAGAGATG 899
DB 840 TCCCTTTCCGATGTTTCAAGAACCGCTTACGTCGCGTACGTCGAGAACCAAGAGATG 900
QY 900 TCGCTCGAGCTGTCAAGACCTCATCGAGTGGCTCAACATCTCGCGGACTCGGCTAT 959
DB 900 TCGCTCGAGCTGTCAAGACCTCATCGAGTGGCTCAACATCTCGCGGACTCGGCTAT 960

```
Qy 960 GTCGTTGACTCTGAACAGTGGATTAACTTCAACCAATGACACACAGTACTACGTCCGCAATC 1019
Db 961 GTCGTTGACTCTGAACAGTGGATTAACTTCAACCAATGACACACAGTACTACGTCCGCAATC 1020
Qy 1020 CGCGTTCTACCTCCAACTAGACGTTCCAGACCCCAAGAGGGGCTTGTTCGACAGTC 1079
Db 1021 CGCGTTCTACCTCCAACTAGACGTTCCAGACCCCAAGAGGGGCTTGTTCGACAGTC 1080
Qy 1080 TCAGACTACCGCTTCTATTAAGGCGATACATGTGAAGCAACATGCCAACACTCGTC 1139
Db 1081 TCAGACTACCGCTTCTATTAAGGCGATACATGTGAAGCAACATGCCAACACTCGTC 1140
Qy 1140 GACCAAGGCTTTTGGATCGGCGGAGTACGCTCTCACCCGACTAGCCTTACCGCAGTAC 1199
Db 1141 GACCAAGGCTTTTGGATCGGCGGAGTACGCTCTCACCCGACTAGCCTTACCGCAGTAC 1200
Qy 1200 GACGTACGAGGCTTACGCTCTGACACTTTGACTTGGCAGACCATCAGCGCGCT 1259
Db 1201 GACGTACGAGGCTTACGCTCTGACACTTTGACTTGGCAGACCATCAGCGCGCT 1260
Qy 1260 GCACTCGGCTTTGTGGGAGGTTTGGCAGAGGTGGCACTGGCGCTGACAGGCACTCCA 1319
Db 1261 GCACTCGGCTTTGTGGGAGGTTTGGCAGAGGTGGCACTGGCGCTGACAGGCACTCCA 1320
Qy 1320 GCCTGGGAGGCACTCTCGGGTGGTACCTCACTGGCGGCCCAACAGGCTACTTTTC 1379
Db 1321 GCCTGGGAGGCACTCTCGGGTGGTACCTCACTGGCGGCCCAACAGGCTACTTTTC 1380
Qy 1380 CCAGCTGGCTTCGTTAGTACGTTCTCCCTGAGGGTTTCGCGCTGAGCGCTACGACCG 1439
Db 1381 CCAGCTGGCTTCGTTAGTACGTTCTCCCTGAGGGTTTCGCGCTGAGCGCTACGACCG 1440
Qy 1440 AACGAGGCTCTTGACCGACTTCGCTTCGCGAGGAGACCGTCACTTTCCGCGCAGTC 1499
Db 1441 AACGAGGCTCTTGACCGACTTCGCTTCGCGAGGAGACCGTCACTTTCCGCGCAGTC 1500
Qy 1500 GCCGTGACGAGGTCGTTGTGACCAACAAACCCCGCGGCGGAGCGCCGCCACCTTC 1559
Db 1501 GCCGTGACGAGGTCGTTGTGACCAACAAACCCCGCGGCGGAGCGCCGCCACCTTC 1560
Qy 1560 ACCGTGAGTGCCCTTCAAAAGCTTACACCAACACCGTGTGTAGGAAACACGCTCTTA 1619
Db 1561 ACCGTGAGTGCCCTTCAAAAGCTTACACCAACACCGTGTGTAGGAAACACGCTCTTA 1620
Qy 1620 GAGACTGACCTCTCTCGTAGGTCGAACTCCCTATGCCACCTGCTGACTTTGGACAG 1679
Db 1621 GAGACTGACCTCTCTCGTAGGTCGAACTCCCTATGCCACCTGCTGACTTTGGACAG 1680
Qy 1680 ACCGTGCGCAACAAACCGGAGATCGAGCAGTCGCTTCTTAAAGAAACACTTTGGCTGCTAT 1739
Db 1681 ACCGTGCGCAACAAACCGGAGATCGAGCAGTCGCTTCTTAAAGAAACACTTTGGCTGCTAT 1740
Qy 1740 TTGTTCACTCAAAATCGAAACCCCGTTTTCAGCTCAGCGAGCGAGCTCTTTGGC 1799
Db 1741 TTGTTCACTCAAAATCGAAACCCCGTTTTCAGCTCAGCGAGCGAGCTCTTTGGC 1800
Qy 1800 GCGTTTCTTCAACAAATCGGTTATGAGCGCACACGCGACCTCCCGGACTACACTGGC 1859
Db 1801 GCGTTTCTTCAACAAATCGGTTATGAGCGCACACGCGACCTCCCGGACTACACTGGC 1860
Qy 1860 ATCCGTGACTCATTCGACAGAAATGTCACCGCTGTGGCGCCACTTCGCTCACTCTCC 1919
Db 1861 ATCCGTGACTCATTCGACAGAAATGTCACCGCTGTGGCGCCACTTCGCTCACTCTCC 1920
Qy 1920 CACTCTGCACTATCGTCACTAAGACCTTACAGGGTTGGGAGCGTCACGAACTCAAC 1979
Db 1921 CACTCTGCACTATCGTCACTAAGACCTTACAGGGTTGGGAGCGTCACGAACTCAAC 1980
Qy 1980 ACGCTTTTCGCGCAATTCGCGCACGCGGCGCTCTCTCAGATGAGGATCTCTGCTC 2039
Db 1981 ACGCTTTTCGCGCAATTCGCGCACGCGGCGCTCTCTCAGATGAGGATCTCTGCTC 2040
```

```
Qy 2040 GCCGACACCTGGCCACCCGCTCTCAGAGTGTCTACCCCGCCTAGCAAACTTCGCGGC 2099
Db 2041 GCCGACACCTGGCCACCCGCTCTCAGAGTGTCTACCCCGCCTAGCAAACTTCGCGGC 2100
Qy 2100 GCGTTTCTGCTTCGCGCGGACATGCTGCTCCCTCGTGAAGTTCGAGGCAACGTC 2159
Db 2101 GCGTTTCTGCTTCGCGCGGACATGCTGCTCCCTCGTGAAGTTCGAGGCAACGTC 2160
Qy 2160 TCATCATCAAGTCCGTTGGCGAGACTGCGCTCGCGCGGCTCAGTCCGCGCTCGCGAAG 2219
Db 2161 TCATCATCAAGTCCGTTGGCGAGACTGCGCTCGCGCGGCTCAGTCCGCGCTCGCGAAG 2220
Qy 2220 CTACCCGAGCTGCTTAATGAGTGTACAGGGAAGATTTCGCGCGCTGTCGCGCGCGCGA 2279
Db 2221 CTACCCGAGCTGCTTAATGAGTGTACAGGGAAGATTTCGCGCGCTGTCGCGCGCGCGA 2280
Qy 2280 GCGCGCGCGCGCGCTGCTGCAATAGTTTCTGCTCGCTCTGCTTCGCGCGTTTCGTA 2339
Db 2281 GCGCGCGCGCGCGCTGCTGCAATAGTTTCTGCTCGCTCTGCTTCGCGCGTTTCGTA 2340
Qy 2340 AACGCGCTGCTTCGCGCACATTTACCGTACCTAAAGACTCTGGTGAAGTCCCGCTCGTTAC 2399
Db 2341 AACGCGCTGCTTCGCGCACATTTACCGTACCTAAAGACTCTGGTGAAGTCCCGCTCGTTAC 2400
Qy 2400 ACAGCGGCTGTCGCGCGGTTGATTCATTCCTCCCAAGCGGCAAGAGGACGTAGTTC 2459
Db 2401 ACAGCGGCTGTCGCGCGGTTGATTCATTCCTCCCAAGCGGCAAGAGGACGTAGTTC 2460
Qy 2460 TGCGTCCCTCGGATACCA 2478
Db 2461 TGCGTCCCTCGGATACCA 2479
```

RESULT 8

```
S43937 LOCUS S43937 2448 bp RNA linear VRL 08-MAY-1993
DEFINITION capsid protein [Nudaurelia capensis omega virus, Genomic RNA, 2448 nt].
ACCESSION S43937
VERSION S43937.1 GI:255203
KEYWORDS Nudaurelia capensis omega virus
SOURCE Nudaurelia capensis omega virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Tetraviridae; Omegatetravirus.
REFERENCE 1 (bases 1 to 2448)
AUTHORS Agrawal,D.K. and Johnson,J.E.
TITLE Sequence and analysis of the capsid protein of Nudaurelia capensis omega virus, an insect virus with T = 4 icosahedral symmetry
JOURNAL Virology 190 (2), 806-814 (1992)
MEDLINE 92391097
PUBMED 1519360
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 112777] from the original journal article.
This sequence comes from Fig. 3.
Map location: segment RNA2.
FEATURES
source 1..2448
/organism="Nudaurelia capensis omega virus"
/mol_type="genomic RNA"
/db_xref="taxon:12541"
367..2301
/gene="capsid protein, Cp"
367..2301
/gene="capsid protein, Cp"
/gene="capsid protein, Cp"
/notes="This sequence comes from Fig. 5; Protein sequence from Fig. 3 does not match conceptual translation; Cp"
/codon_start=1
/product="capsid protein"
/protein_id="AAB23198.1"
/db_xref="GI:255204"
/translation="MDSNSASGKRRSRNRVIAAANTVNVAPKQARGRARRSRANNID NVTAAQELGSLDANVITFTPTNATMPEFRSWARGKLDIDQDSIGWYFKYLDPAT
```

ESARAVGEYSKIPDLVKFSDVDAIREIYNEBECTVSDASIPDLGAQWSLSIISYPMP
 RTAYPAVANVDNKSISIDVNDLJVLNLIASWRDVSOWGFSDPTWFRVIRL
 HPTVDLPDPTGLRLTCSYDLRTVYSITCEANMPTLVDOGFWMIGHVALPTIAATQNA
 VEGSGFERPNTVRPRTAAGVTLTWSHPGSGSAPSGDPAWIPDSITQFQWRHGGFPA
 PTGVIYITIPGYITMQIFDITTNWNGFANFDDVVFQGTGAAGTNATITITAPYTVL
 LTIILATTTSAANINFRNLDAETTAISNRSRSEVPPLPTFGQTAPNNPKIBOTLVKDTL
 GSYLVHSMRNPVQLTPASSFGAISFTNPGFDRNLDPFGFGRISLSDVNMSTAVCH
 FRSLSKCSI VTKYQWEGVTVNTVTFQGFQFAHGLLKNDEILCLADDLATRLTGVVG
 ATDNFAAVALFAANMLTSVLKSEATTSVikelNQATGLANQGLARLPGLLASIPGK
 IAAVRARRDRRRRAAMNN"

BASE COUNT 547 a 715 c 563 g 623 t

Query Match 33.6%; Score 833; DB 14; Length 2448;

Best Local Similarity 62.1%; Pred. No. 2.9e-145;

Matches 1531; Conservative 0; Mismatches 895; Indels 41; Gaps 12;

Qy	3	TTTTCTTTTACCAAGTGTGGTAAATTTAAACAAAGAAAGAAACAGGACCGTAACC	62
Db	1	TTTTCTCTCTTACCAGTGGGTAACAAACAAAGAAAGAAATCGAG---	TGACAA 56
Qy	63	CGGCCCTTACACACCTCGAGTCCGTGACACCGGATTATAGTCGCCCCACACACGCGGC	122
Db	57	CGGCCCTTACACACCTCGAGTCCGTGACACCGGATTATAGTCGCCCCACACACGCGGC	115
Qy	123	CTTTTCCGACCACTCTCGAGAGTCTGTCGGAGTTCGTCCGTGACACCCCGTGGCAGT	182
Db	116	CTTTTCCGACCACTCTCGAGAGTCTGTCGGAGTTCGTCCGTGACACCCCGTGGCAGT	174
Qy	183	CGACAGACGCTTCGGACCACT-AGAACCCTCTCGAGCGAGCGACACACACACACAC	241
Db	175	AAATTTACTGCTTCGGACCACTAAGTTTGTCTCGAAGCGACATACATTC-GTTACACAC	233
Qy	242	CGCCTTACTGCACTACGGAGCGTTGATAGCGCGATTATGAGCGAGCACACCATCG	301
Db	234	TGCTCTTAGCGAGACCTAAACTCGCCTTGATAGACAGCTCATGACAGATCGTTGCC--	GG 290
Qy	302	CCCCTCTCACATTTACCACCGGTATACCCCTTGCCCTTAATACCCCTGAACTGAAG	361
Db	291	CGATCTGTTTCTACTCTCTGCTATTCGTTGCTATTTATCTCTCCGAGC---AAG	347
Qy	362	CAGGATGGAGATGTGGAGTGGCTACAGCGACTTCACAAACCGTCGCGGAACCGTAA	421
Db	348	ACGGCTGGAACTTTTAGAATGGACAGTAACTCAGCCCTCGGCAAGCAGCATCGGCAA	407
Qy	422	CGTTCGGGTACGCGCAACACCGTCACCGTCAATGGTAGAAGAAACCAAC-----	GGCG 475
Db	408	TGTGCGGATTGCGGCAACACCGTCACCGTCAACCGTCTCGCTCCGAAGCAACGCCAAG	CACGCGCGC 467
Qy	476	TGCGACCGGAAGGCAAGTTTCTCCCTCGACAAATTTACCGCTGTCGACAAAGACCTCG	535
Db	468	CGAGCCGCTTACGTGCGCAACAAATGACAAATGTCACTGCGCGAGCTCAGAACTCGG	527
Qy	536	GCAAAAGCCTTGACGCGCAACACCGTCACTTTTCCCGCTAACATCTCTAGCATGCCGGAAT	595
Db	528	CCAACTCTTGAGCGCAAGGTTATTACCTTCCCAACCAATGTTGCAACCATGCGCGAAT	587
Qy	596	CCGGAATTTGGCGCAAGGGAAGATCGACTCGACTCCGATTCCATCGCGTGTACTTCAA	655
Db	588	CCGTAGCTGGGCGCTGGTGAATTTGACATCGACACGAGGACTCCATTGGTTGGTATTCAA	647
Qy	656	GTACCTTGACCCAGCGGTGTACAGAGTCTGCGCGCGCGTCCGCGGAGTACTCGAAGAT	715
Db	648	ATACCTTGACCCCTGTCGCGCACCGAGTTCGCTCGTCCGCTGCGTGGTACTCCAAAT	707
Qy	716	CCCTGACCGGCTCGTCAAGTTCTCGCTCGACGACGAGATAGAGAGATCTATAACAGGA	775
Db	708	ACCTGACCGGCTCGTCAAAATTTCTGTCGACCGCGAGATAGAGAGATATACAATAGGA	767
Qy	776	GTGCCCCGTGCTACTGAGGTGTCGTCGCCCTCGACCGCGCCAGTGGAGGCTCTCGAT	835
Db	768	GTGCCCCCTACCGTCTCTATATACCACTGAGCGCGCCCAATGAGAGCTGTCTAT	827

Qy	836	TTTCTCTCTTCCGATGTTAGAACCGCCTACGTCGCGTAGCGAAACGTCGAGAACAGGA	895
Db	828	CATCTCTACCTATGTTTCCGACTGCTTACTTTGCTGTGCGGAAACGTCGATAAAGGA	887
Qy	896	GATGTCGCTCGAGTGTCAACGACTCATCGAGTGGCTCAACAATCTCGCCGACTGGCG	955
Db	888	GATATCCCTTTCGAGCTACGAAACGATCTTATCGTTTGGCTTAAACAATCTCGCTAGTGGCG	947
Qy	956	TTATGTCGTTGACTCTCAACAGATGGAATAACTTCCACCAATGACACCACTACTACGTCG	1015
Db	948	TGACGTTGTTGATAGCGGACAGTGTTCATCTTTTCGAGCGACCCGACTTGGTTCGTCG	1007
Qy	1016	CATCCGCGTTTACGTCCTCAACCTACGAGCTTCCAGACCCACAGAGGCGCTTGTTCGAC	1075
Db	1008	CATTGCTGTGCTACACCCACATACGATCTTCTGTATCCGACCCGAGGCTTACTTCGTAC	1067
Qy	1076	AGTCTCAGACTACCGCTCCTACTTATAGGGGATTAACATGTGAAGCCACATGCGCAACACT	1135
Db	1068	GTGCTCCGACTTACCGACTTACTTATAAGTCGATTAACCTTGTGAGGCTAACATGCCACGCT	1127
Qy	1136	CGTCGACCAAGGCTTTTGGATCGCGCCGAGTACGCTCTCACCCCGACTAGCCTACCGCA	1195
Db	1128	CGTTGACAGGATTTTGGATCGAGGACATTAAGCTCTTACTCGATCGCCACACACA	1187
Qy	1196	GTACGAGCTCAGCGAGCCTACGCTCTGACACATTTGACCTTCCGCGAGACCATCCAGCGC	1255
Db	1188	GAATGCGTTGAGGTTTGGGTTTGGAGCACCTTTCAACGTCACCCGACCGGCTATCGC	1247
Qy	1256	CGCTGCACTCGGTTTGTGGGCGAGGTTTGGCCACAGGTTGGACTCGGCTGCGAGCAC	1315
Db	1248	AGCGGTTTACCCCTCACTTGGGCTCAATGCGCCCGGTTGGAGTGTCTCGTCCGCTGA	1307
Qy	1316	TCCAGCTGGGAGCAGCATCTCGGTTGGCTACTCACTGCGCGCCACACACGCTACTAC	1375
Db	1308	TCCTGCTTGGATACCCGATCCAC---CAGCGAGTTCAAGTGGGCTCATGAGGTTTGA	1364
Qy	1376	TTTCCGAGCTCGCTCGTTAGCTACTGTTCTCCCTGAGGTTTCCGCTTGAAGCGCTACGA	1435
Db	1365	CGCTCCGACCGGCTCATCTTATCTATCCGCGAGGCTTACACCATGCAATACTTTGA	1424
Qy	1436	CCGAAACGAGCGCTTGGACCGACTTTCGTCGCGAGGAGACACCGTCACTTCCGCGCA	1495
Db	1425	CACTACCAACGAATGGAATGGGTTTGGCAATCCAGATGACGTGGTTACATTCGGGCA	1484
Qy	1496	GGTCGCGTCGACGAGTCTGTTGTGACCAACACCCCGCGCGGCGAGCGGCCCCAC	1555
Db	1485	AACGTGGTGGTCGCGCGGAACTAACGCCACCATTAAGTCACTGACCTACTGTACTCT	1544
Qy	1556	CTTCAACGTCGAGAGTGGCCCTTCAAAAGCTTTACCAACACCGTGTGTTAGAAACAGCT	1615
Db	1545	CACCATTCGCGCACGACCACTTCTGCTGCTTAAACGTCATTAAAC--TTCCGGAATTTGA	1601
Qy	1616	CTTAGAGACTCGACCTCTCTCTGAGGCTCGAACTCCCTATGCACTGCTGCTGCTTGG	1675
Db	1602	CGCTGAAACTACCGCTGCTTCTTAAATCGAGTGAAGTGCCTCTGCGGCTCTTAACTTTGG	1661
Qy	1676	ACAGAGCTGCGCAACCAACCGGAGATCGAGCAGTCTGCTTCTTAAAGAAACACTTGGCTG	1735
Db	1662	TCAAACCTGCGCGCAACCAACCGGAGATGAACAGACTCTGCTCAAGGACACTTGGGTTTC	1721
Qy	1736	CTATTGTCCTCACTCCAAAATCGAAACCCCGTTTTCAGTCTCAGCTCAGCGACGAGCTCTT	1795
Db	1722	ATACCTTGATACAGCAAGATGCGCAACCCCGGTTTTCCAACTCAGCGCTGCTCTTCTT	1781
Qy	1796	TGGCGCGTTCCTTCAACATCGGGTTATGAGGCGACACCGGACTCCCGGACTACAC	1855
Db	1782	TGGCGCGCATAGTTTTCAGAAATCCCGGTTTGAACCGCAACCTCGATCTCTCTGTTGG	1841
Qy	1856	TGGCATCGCTCACTCATTCGACACCAACATGTCACCGCTGTCGCGGCTTCTCTGTTGCT	1915
Db	1842	CGGATACGTCGCTCTCTTGTGATGTCACATGTCACCGCTGTCGCTGTCCTTCTT	1901
Qy	1916	CTCCCACTCTCTGAGTATCTGTCATTAAGACCTAACAGGTTGGGAAGCGGCTCACGAACGT	1975

1902	DB	GTCTAAGTCATGTTCTATTGCTCACTAAGACGTACCAAGGTTGGGAAGGTTTACTAATGT	1961
1976	QY	CAACACGCCCTTTTCGGCCAAATTCGCGACGCGGCCCTCTCAAGAATGAGGAGATCCTCTG	2035
1962	DB	TAAACACCCCTTTTGGTCAATTTGCTCACAGCGGGCTGCTCAAGANTGACAGATTCCTTTG	2021
2036	QY	CCTCGCGACGACTGGCCACCCGCTCTCACAGGTGTCTACCCGCCACTGCACAACTTCG	2095
2022	DB	CCITGCAGACGACTTAGCTACCCGCTCTTACAGCGGTTATGGTGCACGATAATTTTGC	2081
2096	QY	GGCGCGCGTTTCTGCCCTTCGCGCGAACAATGCTGTCTCGTGCTGAAFTCGAGGCAAC	2155
2082	DB	TGCTGCGGTCCTCGCTTTTCGCGCAAAATATGCTCACTTCGCTTGAAGTCGGAAGCTAC	2141
2156	QY	GTCTTCATCATCAAGTCGTTGGGAGACTGCGGTGCGCGCGGCTCAGTCCGGCCCTCG	2215
2142	DB	CACCTTCAGTTATTAAAGAGAGCTCGGCAATCAAGTACTGGGACTTGGCAACAGCGGCGCTGC	2201
2216	QY	GAAGCTACCCGGAGCTGCTAATGAGTGTACCAGAGGAAGATTGCCGCGTGTCCGCGCGG	2275
2202	DB	CCGTTTACCGGGTCTCTTCGCTCGATACCGGGCAAAATGTGTGACAGTGTCCGCGCTAG	2261
2276	QY	CCGAGCGCGCGCGCGCGCTCG--TGCCAAATTAGTTTGTCTCGCTCCTGTTTTCGCGGTT	2333
2262	DB	ACGTGACCGTGCAGTGCTGCAGATGACAAATTAACACGTGGGCACATTTCCCGT	2321
2334	QY	TCGTAAACCGCGTGGTCCGCGACA-----TTACGCGTACCCCTAAAGACTCT	2380
2322	DB	TTTAGGAAACGGGTGGTACTGATCAACCATGAAACCTGGCTATCTGATCTAATAAACTCT	2381
2381	QY	GGTGAGTCCCGCTGTTTACACGAGGCTTCGCGCGGTTGATTCGATTCGATCCCAAGCGGCA	2440
2382	DB	GGTGAGTTCCTGTGCACACACGACAGTTTTCGTCCGGTTCGATTCGAATCCCAAGCAGCG	2441
2441	QY	AGAAGGA	2447
2442	DB	CGGGCA	2448

```

RESULT 9
AF548354
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
2593 bp RNA linear VRL 27-MAR-2003
Providence virus capsid protein precursor (CP) gene, complete cds.
AF548354
AF548354
AF548354.1 GI:29293088
Providence virus
Providence virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Tetraviridae;
unclassified Tetraviridae.
1 (bases 1 to 2593)
Pringle,F.M., Johnson,K.N., Goodman,C.L., McIntosh,A.H. and
Ball,L.A.
Providence virus: a new member of the tetraviridae that infects
cultured insect cells
Virology 306 (2), 359-370 (2003)
22529782
12642108
2 (bases 1 to 2593)
Pringle,F.M., Johnson,K.N. and Ball,L.A.
Direct Submission
Submitted (23-SEP-2002) Microbiology, University of Alabama at
Birmingham, 845 19th Street South, BBRB 373/17, Birmingham, AL
35294, USA
Location/Qualifiers
1..2593
/organism="Providence virus"
/virus
/mol_type="genomic RNA"
/specific_host="Helicoverpa zea midgut cell line"
/db_xref="taxon:213633"
/note="also infects Helicoverpa zea fat body cell line"

```

gene	CDS	mat_j	mat_j	mat_j	BASE COUNT ORIGIN

[illegible]

Db 1220 TACGCAATGGCTGAGATCGCGGATAGGACTG---ACAGTGTATCCAGTTTAGGAAGTG 1276
Qy 1098 TATAAGCGATAAATGTAAGACCAATGATGCAACATCGTCGACCAAGGCTTTGGATC 1157
Db 1277 TATAAGGAAATTAATCTTTCAGATTCAATGCCCCCACTTAATGATCAAGGTTGGTGGTG 1336
Qy 1158 GCGCGCAGTAGTCTCACCCTGAGTACGCTACCGAGTAGTACGAGTCAAGGAGGCTTAC 1217
Db 1337 GGTGACACATCCCTGTAAAGCC-ACATCAGAACAACATCCCTGCTGAGAGGTTCTC 1395
Qy 1218 GCTGTGACATTTGACCTTCGCCAGACCATCCAGCGCGCTGCACTCGCGTTTGTGTGG 1277
Db 1396 AGCTGGTTCAATGACAGTTTCGGCTTCAATGCTATCTTCCAAACCATTAACACTGTGCGC 1455
Qy 1278 GCAGGTTTGCACAGGTTGGCACTGCGCTCGAGGCTCGAGGCTGGAGCAGGCATCC 1337
Db 1456 CAGAATTGTTGGAGCATAAAGCGCTCCCGGTAGCTACAGTAGCTTTGACTACTGGAAC 1515
Qy 1338 TCGGTGGCTACCTCACCTGCGGCACAAACGGTACTACTTTCCAGCTGGCTCGGTAGC 1397
Db 1516 TGGAGGCACTAACAT-----ACAAGCGAAAGTTCTTCAGTGTGGAGATCGATGG 1566
Qy 1398 TACGTTCTCCCTGAGGTTTCGCCCTTGAGCGCTAAGACCGCAACGAGGCTCTTTGACC 1457
Db 1567 TAACTCACTCAGTGTGAGCTTTTACCGCCCACTTTCCATATTTAGCGAAGGTGAACC 1626
Qy 1458 GACTTCGCTTCGAGGAGACACGCTCACTTTCCGCGAGGTGCGCGTCAAGAGGTCGTT 1517
Db 1627 -TTTTCAGAGGAGGAGACAACTTCGTTTCAATGAGCACTATCACGGCAGATACT 1684
Qy 1518 GTGACCAACAAACCGCGCGGCGGAGCGCCCACTTCACCGTGAGAGTGCCTT 1577
Db 1685 GTTGTACTCAGTTAGCAGTTCTTTGACTGGTTCATCCGTAATTTGTGAGAGTGTGAGC 1744
Qy 1578 TCAACGCTTACCAACACCGTGTTTAGGAACAGCTCTTTAGAGACTCGACCTCTCTCT 1637
Db 1745 AAAGGCTGTGGCTCAGTATCACTCTCTGCTGCTAGTATAGATACCGAGCTGTAAAT 1804
Qy 1638 CGTAGGCTGAACCTCCCTATGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1697
Db 1805 C---GGTTGAGTATGAGATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1861
Qy 1698 AAGATCGAGAGTCTCTTTAAAGAAACACTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1757
Db 1862 AATACGAGAGTCTCTATGTAAGAAATCCGAGGAGGATACATTTGTTCACTATAAGATG 1921
Qy 1758 CGAAACCCGTTTTCAGCTCAAGCAGCAGCTCTTTTGGCGCGTCTTCTTCAACAAAT 1817
Db 1922 AACACCTCTGTTTGTAGATGACTGGGAAGAAACTTCGGTGGATTCCAATTCACATAC 1981
Qy 1818 CCGGTTATAGCGACACAGGACCTCCGAGTACACTGCGATCCGCTGCTGCTGCTGCTGCT 1877
Db 1982 CCAGGATATGATCCGGAATAAATGCACTAGTCTTAAGAGGCTTTGTAGACACCTTCGAA 2041
Qy 1878 CAGAACATGTCACCGCTGCGCCCACTTCGCTCACTCTCCACTCTGCTGCTGCTGCTGCT 1937
Db 2042 AACAACTTCAAGTGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2101
Qy 1938 ACTAAGACTTACAGGTTTGGGAAGCGCTACGAAAGTCAACACGCTTTTCGGCAATTC 1997
Db 2102 TGTAAACATATGATGATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2161
Qy 1998 GCGACGCGGCTCTCTCAAGATGAGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2057
Db 2162 GCCACACTGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2221
Qy 2058 CGTCTCAGAGTGTCTACCGCCCACTGACAACTTCGCGCGCGCGCTTCTGCTGCTGCTGCT 2117
Db 2222 GAACCTAACCGGAGTTTACAGGCTGATGATTAATTTGCGGCACTGTGAGCGCTTTGGCC 2281
Qy 2118 GCGAACATGCTGCTCCGCTGCTGAGTGGAGGCAACGCTCTCCATCATCAAGTCCGCTT 2177

Db 2282 AGTATCGGCTGGACATTTAGGTAAAGCTCGGCAACCGTCAAGTTATTAAAGGGATC 2341
Qy 2178 GCGAGACTCCGTCGCGCGGCTCAGTCGCGGCTCGCAAGCTACCCGAGCTCTCTAA-- 2235
Db 2342 GCACAACAAGGCTCGCGCTGTACAGCTAACCTGTATCTTAGAAGCGCGGTGAAA 2401
Qy 2236 -----TGAGTGTACAGGAAGATTGCGCGGCTGTCGCGCGCGCGCGCGCGCGC 2288
Db 2402 GCCATTGTTGAGTGTGAGCAAGGCTAGTCTGGGTGCAATTAAAGGCACGCTCGCGCGAGCGC 2461
Qy 2289 CGCGCGCTCGTCCCAATTAGTTGCTCGCTCCCTGT 2324
Db 2462 CGCGCTAGACGAGCAAGTAAAGTGGCCACTTCGCT 2497

RESULT 10
AR125092 LOCUS 5312 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 39 from patent US 6177075.
ACCESSION AR125092
VERSION AR125092.1 GI:14111154
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5312)
AUTHORS Christian,P.Daniel., Gordon,K.Hienrich.Julius. and Hanzlik,T.Nelson.
TITLE Insect viruses and their uses in protecting plants
JOURNAL Patent: US 6177075-A 39 23-JAN-2001;
FEATURES Location/Qualifiers
 1..5312
 /organism="unknown"
BASE COUNT 1133 a 1740 c 1432 g 1007 t
ORIGIN

Query Match 2.9%; Score 71.2; DB 6; Length 5312;
Best Local Similarity 78.7%; Pred. No. 0.003;
Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 2371 TAAAGACTCTGTTGAGTCCCGCTGTTACAGCAGGCTCTGCGGGTTCGATTCCTC 2430
Db 5205 TAAAGACTCTGTTGAGTCCCAAGCTTACTGTTGAGTCTGCTCGGTTTCGATTC 5264
Qy 2431 CCAAGCGCAAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2478
Db 5265 CCAAGCGCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5312

RESULT 11
AR125093 LOCUS 5312 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 41 from patent US 6177075.
ACCESSION AR125093
VERSION AR125093.1 GI:14111155
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5312)
AUTHORS Christian,P.Daniel., Gordon,K.Hienrich.Julius. and Hanzlik,T.Nelson.
TITLE Insect viruses and their uses in protecting plants
JOURNAL Patent: US 6177075-A 41 23-JAN-2001;
FEATURES Location/Qualifiers
 1..5312
 /organism="unknown"
BASE COUNT 1133 a 1740 c 1432 g 1007 t
ORIGIN

Query Match 2.9%; Score 71.2; DB 6; Length 5312;
Best Local Similarity 78.7%; Pred. No. 0.003;
Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 2371 TAAAGACTCTGGTGGAGTCCCGTGTACACGACGGGTCTGCCGGTTCGATTCATTC 2430
Db 5205 TAAAGACTCTGGTGGAGTCCCGTGTACACGACGGGTCTGCCGGTTCGATTCATTC 5264
Qy 2431 CCAGCGGCAAGAGGAGGTAGTCTGCTCGTCCCTCGGATACCA 2478
Db 5265 CCAAGCAGCAAGAGGTCGCAACTAGTACGCGCCCTCGGATACCA 5312

RESULT 12
AR125094 LOCUS 5312 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 43 from patent US 6177075.
ACCESSION AR125094
VERSION AR125094.1 GI:14111156
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5312)
AUTHORS Christian, P. Daniel., Gordon, K. Hienrich, Julius. and Hanzlik, T. Nelson.
TITLE Insect viruses and their uses in protecting plants
JOURNAL Patent: US 6177075-A 43 23-JAN-2001;
FEATURES Location/Qualifiers
source 1. 5312
/organism="unknown"
BASE COUNT 1133 a 1740 c 1432 g 1007 t
ORIGIN

Query Match 2.9%; Score 71.2; DB 6; Length 5312;
Best Local Similarity 78.7%; Pred. No. 0.003;
Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Qy 2371 TAAAGACTCTGGTGGAGTCCCGTGTACACGACGGGTCTGCCGGTTCGATTCATTC 2430
Db 5205 TAAAGACTCTGGTGGAGTCCCGTGTACACGACGGGTCTGCCGGTTCGATTCATTC 5264
Qy 2431 CCAGCGGCAAGAGGAGGTAGTCTGCTCGTCCCTCGGATACCA 2478
Db 5265 CCAAGCAGCAAGGTCGCAACTAGTACGCGCCCTCGGATACCA 5312

RESULT 13
HAU18246 LOCUS 5312 bp. RNA linear VRL 29-DEC-1998
DEFINITION Helicoverpa armigera stunt virus RNA-dependent RNA polymerase, complete cds.
ACCESSION U18246
VERSION U18246.1 GI:603222
KEYWORDS
SOURCE Helicoverpa armigera stunt virus
ORGANISM Helicoverpa armigera stunt virus
REFERENCE 1 (bases 1 to 5312)
AUTHORS Gordon, K.H., Johnson, K.N. and Hanzlik, T.N.
TITLE The larger genomic RNA of Helicoverpa armigera stunt tetravirus encodes the viral RNA polymerase and has a novel 3'-terminal tRNA-like structure
JOURNAL Virology 208 (1), 84-98 (1995)
MEDLINE 21820274
PUBMED 11831734
REFERENCE 2 (bases 1 to 5312)
AUTHORS Gordon, K.H.J.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-1994) Karl H.J. Gordon, CSIRO Division of Entomology, Biotechnology Section, Canberra, Act, Australia, 2601
FEATURES Location/Qualifiers
source 1. 5312
/organism="Helicoverpa armigera stunt virus"
/mol_type="genomic RNA"

CDS

/specific_host="Helicoverpa armigera"
/db_xref="taxon:37206"
37..5151
/codon_start=1
/product="RNA-dependent RNA polymerase"
/protein_id="AAC98529.1"
/db_xref="GI:603223"
/translation="MYAKATDVAVVAAADVAANVLQRAVLDKDFAPPLKALETLHR
NAPVADYHGCTGTGRDGRSHITALESRVATGPEFKADASLTASRTFCVDDGV
GSCAFKSRGIANHSLYDVTLELANAFENHGLHVMRAFMHPEELLMDNVNAELG
YRPHVTEEMAVKDCAFQGGDLRHPPELDIFNESQERRTERLAARSGSERRAVIFSG
DDWGDAYLHDFHTLWYLLVRYNPTFPFSLHIEVQRHGSSELRLITRAPGCDML
AVPRTSQGLCRIPFIYADASGTEHKTLITLSQHKVNLMLNFERPELKVAKTDVL
MSFARALRAIVAVSEVTESSWNISPADLVTVVSVLYLHIIERRRAAGVPAATKADV
FLPSRVPFARVHDFRDELEVLREAGCVNERPVPSTPEVPEPQGFADLWMTAASLPE
YRATLOAGLNTDVOKLTLENAKLTIDGLTSPVRGLEMYEGPGSGKGTGLIAALE
AAGKALYVAPTRELAMDRRIKPPSASTQHVALAILRATAEGAPFATVIDEFC
MFFLVYVAIVHASPSSRIVLGDVHQIFDQGTSAANPLVDRVQVQRRRTFNOT
KRCPADVATTPFQSLYPGCTTSGCVASISHVADPNRSQAOTCLCTQSEKGRHGAER
GAMTVEHAQRTFASVILHYNGSTAEQKLLAEKSHLIVGITRHTNHLIIRDPTGDIER
OLNHSKAEVFTDIPAPLEITTVKPESEVORNEVMAIIPQSPTHCAIHLKRNFGD
OPDGCVALAKTGYEVGGRAKTNVLAEPDAPKPHRAFOEGVQVWVKNVSNKHOA
LQTLRLRYKRSADLPLHEAKEDVKMLNSLDHNDWTVTDEDRAVFEQTKETQR
GGTVEDLLEDDPYIRIDIDFMKTQKQVSPKINTKVGQGIHAKSLSNPLVLAWLR
ILSEILRTGRTVYRNGLPDEEAMLMLEAKINQVPHATFVSADWTFDFAHNNTSEL
LFAALLERIGTPAAAVKGLDLSLGGSHYLRFDASRLHMGERYKTKHLKVEQKIVPYIG
AVMLTLFRGVKFAAFKGLDLSLGGSHYLRFDASRLHMGERYKTKHLKVEQKIVPYIG
LLVSAEQVLDLPVRSALKIFRCYTSLLYKYVEAVRDITKGSMDARYHSLILCHMSA
CYNYPESAAYIIDAVVREGRDFFPEQLRVRAHVQADAYSSITPANVRASCLDH
VFEPROAAPAGFVATCAKPTSPSTAKAGVSAATTSVAITGATPAPPSPMDAPANSF
SELLTPTPTSSSAVIVIFGLLYIVKVAQWHRKRTDLNSRKPSPQSDROSRSE
CLDRSGERTGSSLTAPTAPSPFSFSEARLATPTVAAATSPSATSCATDQVAART
TPDPAPLGSQSAKAVSKVPYPTTARWKEVTPHAWKGVTDGRPEVREDEPETAAVQ
ALISGRYPQTKLSSDASKGYSTKCSQSTSPAPSADYQARDQCVTRVCRAAEWA
RSCIHPLASSAASADLKRIRSTSDSPVDVKISKA"
misc_structure 5190..5312
/note="3'-terminal tRNA-like structure can be formed by
this region; lacks pseudo-knot in aminoacyl acceptor stem;
has valine anticodon"
BASE COUNT 1133 a 1739 c 1433 g 1007 t
ORIGIN

Query Match 2.9%; Score 71.2; DB 14; Length 5312;
Best Local Similarity 78.7%; Pred. No. 0.003;
Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Qy 2371 TAAAGACTCTGGTGGAGTCCCGTGTACACGACGGGTCTGCCGGTTCGATTCATTC 2430
Db 5205 TAAAGACTCTGGTGGAGTCCCGTGTACACGACGGGTCTGCCGGTTCGATTCATTC 5264
Qy 2431 CCAGCGGCAAGAGGAGGTAGTCTGCTCGTCCCTCGGATACCA 2478
Db 5265 CCAAGCAGCAAGGTCGCAACTAGTACGCGCCCTCGGATACCA 5312

RESULT 14
AR125095 LOCUS 5368 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 45 from patent US 6177075.
ACCESSION AR125095
VERSION AR125095.1 GI:14111157
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5368)
AUTHORS Christian, P. Daniel., Gordon, K. Hienrich, Julius. and Hanzlik, T. Nelson.
TITLE Insect viruses and their uses in protecting plants
JOURNAL Patent: US 6177075-A 45 23-JAN-2001;
FEATURES Location/Qualifiers

```
source 1. .5368
/organism="unknown"
BASE COUNT 1153 a 1747 c 1454 g 1014 t
ORIGIN
Query Match 2.9%; Score 71.2; DB 6; Length 5368;
Best Local Similarity 78.7%; Pred. No. 0.003; ~
Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Qy 2371 TAAAGACTCTGAGTCCCGTCTGTTACAGAGGGTCTCCGCGGTTCGATTCATTCC 2430
Dy 5261 TAAAGACTCTGAGTCCCGTCTGTTACAGAGGGTCTCCGCGGTTCGATTCATTCC 5320
Qy 2431 CCAAGCGCAAGAGGAGTGTAGTCTGCTGCTCCCTCGGATACCA 2478
Dy 5321 CCAAGCAAGAGGAGTGTAGTCTGCTGCTCCCTCGGATACCA 5368

RESULT 15
AF429315/c 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 125020)
Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Strevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)
11594876
2 (bases 1 to 125020)
Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
Direct Submission
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
Location/Qualifiers
1..125020
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and WI-12410"
/note="isolated from a patient with Huntington's
Disease-Like 2 (HDL2)"
complement(35581..35746)
/rpt_type=tandem
/rpt_unit=ctg
complement(<36507..>36887)
/gene="JPH3"
/note="synonym: JP3"
complement(<36507..>36887)
/gene="JPH3"
/product="JPH3"
/product="JPH3"
complement(<36507..>36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="JPH3"
/protein_id="AA14094.1"
/db_xref="GI:17646245"
translation="MSSGGRPNFDGSGYCGWEDGKAHGHGVTGPKQGQGYTGSMG
HGFVLGYVTPWNGNTYGGTWAQKRGHIGLESKRWYKGEWTHGFKRGYVRECA
NGAKYEGTWSNGLQDGYGTETYSBG"
BASE COUNT 29056 a 30696 g 28283 t 4254 others
```

```
ORIGIN
Query Match 2.5%; Score 62; DB 9; Length 125020;
Best Local Similarity 10.8%; Pred. No. 0.11;
Matches 162; Conservative 642; Mismatches 671; Indels 24; Gaps 4;
Qy 132 CCATCTCTCGAGAGTCTGGAGTTTCGTCGTCGACACCCCGTTGCGAGTCGACGACG 191
Dy 52158 MSYCRSGRGRGKMTGYRGMWRSMCWSCCYCKWCMACMMAKSCMCKSCYMYMSMY 52099
Qy 192 CTTCCGAGACCACTAGAACCTCTCTGAGCGAGGACACACAGCAGCAGCAGCAGCAG 251
Dy 52098 RGSNSTCCCMKYRSKYSYKSGWSYMSYKSCASCMGWSYKSRKSGCWSAT 52039
Qy 252 GCACCTAGCAGCGTTGATAGCGCGATTTATGAGCGAGCAGCAGCAGCAGCAGCAG 311
Dy 52038 SYKMYGSKTRSTGGKRSAKMC-----WKWSSNSRSMCMYAWKMGKMYWYRYS 51984
Qy 312 CACATTACCAACCCGTTTACACCTTGCCCTTAATACCCCTGAACTGAAAGCAGGATCG 371
Dy 51983 RKRYSCYRSGSKYKSRWAKSMWRGSMRSMYVWYRMCWKGRCRRGAGSAARSYKKGAM 51924
Qy 372 GATGCTGAGTGGGTGACAGGACCTCAACCGTCGCGGAAACCCGTAAGCTTCGGGTC 431
Dy 51923 KSKGMRGCGSTYYGKMGCRGSKSMYTGSMKCAATGKMSWMMCMGRCAMWSKGGSYRG 51864
Qy 432 AGCGCAACACCGTCAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 491
Dy 51863 WGCWMSYWCASYSRWRGSAWSYAGGTVTSYKSGWKGKMYKGRARSGMSRWKRS 51804
Qy 492 GTTCTCCCTGACAAATTTACCGCTGTCGCAACAGCCTCGCGCAAGACCTTGACGCC 551
Dy 51803 WKKSWRKGSRASARWKSYSASRRKSCWRMMRGCSYMTGRSSMGTMKSRKSGRTSMGS 51744
Qy 552 AACACCTGCTATTTCCCGCTAACATCTTAGCATGTCGCCGAATTCGGAATTCGGCAAG 611
Dy 51743 KAGTYSARAAAYRSCSKRRRGKKSKSGGGRGKKWMASTYSKRCCKAKSTGRRRG 51684
Qy 612 GGAAAGATCGACCTCGACTCCGATTCATCCATCGCTGCTGCTGCTGCTGCTGCTGCTG 671
Dy 51683 TKSSSCWTKRGRGRTKAAAWKACCCWYRCTGCMWAKTYATYATYKSWKSRWTTGWR 51624
Qy 672 GGTGCTACAGAGTCTGCGCGCGCTGCGGAGTACTCGAAGATTCCTCGACGGCTCGTC 731
Dy 51623 MGTERRAMWYWKYRARMKCAKTRSMWMAKSKYSYARCKCSMWSYKYSKRSR 51564
Qy 732 AAGTCTCTCGACGAGAGATAAGAGATCTATTAACAGAGAGTCCCGCTCGTCACT 791
Dy 51563 CKGKWKMAKCMRRYMYKWKGGAGKCCSAGMTGRTGSRWSACCMWMSKKSASARYAY 51504
Qy 792 GACGTGCTCCCTCCCTCGACGCGCGCTGAGGCTCTCGATTTCTCTCTTTCCGATG 851
Dy 51503 ASTGAAAMSGCTCYMCKMCTRRMMATACAWRRRAWMMSCMRKCTSKWSKMGRRMYC 51444
Qy 852 TTCAGAACCCGCTACGTCGCGCTAGCGAAGCTCGAGAACAAAGAGATGTCGTCGACGTT 911
Dy 51443 WSWARWYAGSYMSKCSRWGKYGRSR-SGMSWRTRGYSYARCYMSKSGAGGAKC 51385
Qy 912 GTCAACACCTCATCGAGTGGCTCAACAATCTGCCGAGCTGGCGGTATGCTGCTGCTCT 971
Dy 51384 KGAGSYAKGSAGRWGSKYKNGAWYSCRMSASKSRMTCYMSCCTCKGCKYMYWSWKY 51325
Qy 972 GAACAGTGGATTAACTTCAACCAATGACACCACTGCTACCTCGCATCCGCTTCTACGT 1031
Dy 51324 AGMCWYRWCTCTNARAWRRARMAARRARWAGARRAWMSWMMWMMHHAAMARM 51265
Qy 1032 CCAACCTACGAGTTCAGACCCCAAGAGGCT-----TGTTGCGACAGTCTCAGACTA 1087
Dy 51264 WAAAKCWWYSYRKYCWMSWMAWAAMMMCTAAWYKKTMRMRWSKYSSSYAVAW 51205
Qy 1088 CGGCTCTACTATAAGCGGATTAACATGTGAAGCAACATGCAACATCTGTCGACCAAGG 1147
Dy 51204 YATMCWYCYSMYAAARMGTYSGGKRWASTSCMCCCRGCKMKKGMWCSYTYTGMS 51145
```

```
Qy 1148 CTTTGGATCGGCGGCGAGTACGCTCTACCCCGACTAGCCTACCGCAGTACGACGTCAG 1207
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
51144 GGRSRWNTSSMACWYTKGSRRTKSMATSGCMRMGAMRSGKGMRYKWSCRYKGMRWGRW 51085

Qy 1208 CGAGGCTACGCTCTGACACTTTGACTTCGCGCAGACCATCCAGCGCGCTGCATCGC 1267
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
51084 SGYKCSGYMGMGYSGRTSKYCSYTGVCSCSKAMCKSKTSKSCCTSKYSKSMGR 51025

Qy 1268 GTTGTGTGGCGAGGTTTGACACAGGCTGCGCTGCGGCACTCCAGCCTGGGA 1327
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
51024 YSSYCTCWGRMCKGMKSSCCASRSGCKSYKMAGSWRCRGRGMSCWYWSMKSYSMASY 50965

Qy 1328 GCAGGCATCCTCGGCTGCTACTACCTCGCGCCACACGGTACTACTTTTC----- 1380
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
50964 KSSARGRMCKYKTSCKMYYSMHBSHDSGYKSYRKYBYWTCKMKDSMCWSVDSHM 50905

Qy 1381 -----CAGCTGGCTCGTTAGTACGTTCTCCCTGAGGTTTCGCGCTGAGCGCTAC 1433
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
50904 MRDMYDKMRKVSSGTRGYHBHSAWCKMSRSDSGYSGMVHMKSYWKSSASKHMYYS 50845

Qy 1434 GACCCGAACGACGCTCTTGACCGACTTCGCTTCCGAGGAGACACGTCACCTTCCGG 1493
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
50844 GMKSSVYBDCMSNYTSBSKSVBRBRWSGSYBRYRCAVMRYBHRSTRSVWYGGKRS 50785

Qy 1494 CAGTTCGCGCTCGACGAGTCTGTTGTGACCAACCCCGCGCGCGCGCGCCCC 1553
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
50784 CWVRDSVSRRRSVHTSMSSWRMCRGARMWAKSCAGKMSCVMTGCMVHSRCSGSMKWCSC 50725

Qy 1554 ACCTTACCGTGAGTGGCCCTTAAACGCTTACACCAACACCGTGTGTTAGAACAC 1612
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
50724 DDCCHRABADAHGVCRCBWKBCADBCYKRDVSCWSAHVWVWVBBHBSKSRBGWRHMM 50666
```

Search completed: January 12, 2004, 21:07:06
Job time : 8980 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 17:28:25 ; Search time 660 Seconds
(without alignments)
10135.163 Million cell updates/sec

Title: US-09-991-262-47

Perfect score: 2478

Sequence: 1 GTTTTCTTCTTTTACCAAG.....CTGCGTCCCTCGGATACCA 2478

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*

- 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2478	100.0	2478	19	AAI99117
2	2468.4	99.6	2478	15	AAQ58523
3	158	6.4	172	15	AAQ58525
4	71.2	2.9	164	15	AAQ58524
5	71.2	2.9	5312	15	AAQ58522
6	49	2.0	985	24	ABQ44816
7	49	2.0	985	24	ABQ44817
8	49	2.0	4403765	22	AAI99683

9	49	2.0	4411529	22	AAI99682	Mycobacterium tube
c 10	47.6	1.9	10732	21	AAAI0594	Gene encoding a su
c 11	46	1.9	4403765	22	AAI99683	Mycobacterium tube
c 12	46	1.9	4411529	22	AAI99682	Mycobacterium tube
13	44.6	1.8	3285	25	ACC44572	Alpha-amylase/glc
14	44.4	1.8	1362	25	ABX56041	M. echinospora cal
15	44.2	1.8	1221	23	AA54074	Pseudomonas aerugi
16	43.8	1.8	815	19	AAV64513	M. tuberculosis im
17	43.8	1.8	815	19	AAV44404	Mycobacterium tube
18	43.8	1.8	815	20	AAZ19314	M. tuberculosis an
19	43.8	1.8	815	20	AAZ19102	M. tuberculosis re
20	43.4	1.8	774	22	AAZ16362	Human SBhPRO221 ge
21	43.4	1.8	774	22	ABQ86147	Novel human gene.
22	43.4	1.8	1314	14	AAQ43972	PEP PM DNA. Strep
c 23	43.4	1.8	349980	24	ABQ81844	Bifidobacterium lo
24	43	1.7	6607	24	ABK91456	Modified HIV prote
25	43	1.7	6784	24	ABK91460	Modified HIV prote
26	43	1.7	6787	24	ABK91455	Modified HIV prote
27	43	1.7	6856	24	ABK91462	Modified HIV prote
28	43	1.7	6859	24	ABK91458	Modified HIV prote
29	43	1.7	6965	24	ABK91459	Modified HIV prote
30	43	1.7	6970	24	ABK91454	Modified HIV prote
31	43	1.7	7003	24	ABK91521	Modified HIV prote
32	43	1.7	7037	24	ABK91461	Modified HIV prote
33	43	1.7	7042	24	ABK91457	Modified HIV prote
34	43	1.7	7181	24	ABK91518	Modified HIV prote
35	43	1.7	7186	24	ABK91520	Modified HIV prote
36	43	1.7	7186	24	ABK91523	Modified HIV prote
37	43	1.7	7315	24	ABK91525	Modified HIV prote
38	43	1.7	7351	24	ABK91452	Modified HIV prote
39	43	1.7	7493	24	ABK91519	Modified HIV prote
40	43	1.7	7498	24	ABK91524	Modified HIV prote
41	43	1.7	7498	24	ABK91527	Modified HIV prote
42	43	1.7	7547	24	ABK91449	Modified HIV prote
43	43	1.7	7558	24	ABK91451	Modified HIV prote
44	43	1.7	8008	24	ABK91453	Modified HIV prote
45	43	1.7	8009	24	ABK91450	Modified HIV prote

ALIGNMENTS

RESULT 1	
AAI99117	
ID	AAI99117 standard; cdna; 2478 BP.
XX	
AC	AAI99117;
XX	
DT	08-JUN-1998 (first entry)
XX	
DE	Helicoverpa armigera RNA2 encoding coat proteins p71 and p17.
XX	
KW	Vaccine; coat protein; p71; p17; insecticide; Ig-like domain; ds.
XX	
OS	Helicoverpa armigera.
XX	
PH	Key
FT	Location/Qualifiers
FT	283..756
FT	/*tag= a
FT	/product= p17 coat protein
FT	366..2309
FT	/*tag= b
FT	/product= p71 coat protein
XX	
PN	WO9746666-A1.
XX	
PD	11-DEC-1997.
XX	
PF	02-JUN-1997; 97WO-AU00349.
XX	
PR	31-MAY-1996; 96AU-0000234.
XX	
PA	(CSIR) COMMONWEALTH SCI & IND RES ORG.

XX	PI	Gordon KH, Hanzlik TN;	661	TTGACCCAGCGGTGCTACAGAGTCTGCGCGCGCGTCCGCGAGTACTCGAAGATCCCTG	720
XX	PI	WPI; 1998-042175/04.	721	ACGGCCCTCGTCAAGTTCTCCCGTCGACGCGAGAGATCTTATAACGAGGAGTCC	780
DR	DR	P-PSDB; AAW34534, AAW34535.	721	ACGGCTCTGCTCAAGTTCTCCCGTCGACGCGAGAGATCTTATAACGAGGAGTCC	780
XX	PT	Modified small RNA viruses and virus-like particles - have altered	781	CGCTGCTACTGACGTGCTCGTCCCTCCGACGCGCGCGAGTGCAGGCTCTCGATTTTCT	840
PT	PT	or substituted Ig-like domains to modify host cell tropism, useful	781	CGCTGCTACTGACGTGCTCGTCCCTCCGACGCGCGCGAGTGCAGGCTCTCGATTTTCT	840
XX	PS	as insecticides and in medicinal applications	841	CTTTCCGATGTTCAAGACCGCTACGTCGCGTACGAACTCGAGAACAGAGATGT	900
XX	PS	Disclosure; Figure 1; 41pp; English.	841	CTTTCCGATGTTCAAGACCGCTACGTCGCGTACGAACTCGAGAACAGAGATGT	900
CC	CC	The sequence is that of a cDNA encoding RNA2 of the viral genome which	901	CGCTGACGCTTGTCAACGACCTCATCGAGTGGCTCAACAATCTCGCCGACTGGGTTATG	960
CC	CC	putatively codes for coat proteins p71 and p17. The p71 coat protein	901	CGCTGACGCTTGTCAACGACCTCATCGAGTGGCTCAACAATCTCGCCGACTGGGTTATG	960
CC	CC	includes an Ig-like domain which can be used in the production of	961	TGCTTGACTCTGAAACAGTGGATTAATCTTCAACAATGACACACGTACTACGTCGCAATCC	1020
CC	CC	virus-like particles (VLP). The VLPS can be used in vaccines where the	961	TGCTTGACTCTGAAACAGTGGATTAATCTTCAACAATGACACACGTACTACGTCGCAATCC	1020
CC	CC	Ig-like domain has been altered so that the VLP presents a surface	1021	GGCTTCTAGTCCAACTACGAGTTCAGACCCCGACAGAGGGCTTGTTCGACAGTCT	1080
CC	CC	located antigen which is used to elicit an immune response in a host	1021	GGCTTCTAGTCCAACTACGAGTTCAGACCCCGACAGAGGGCTTGTTCGACAGTCT	1080
CC	CC	organism. They can also be used controlling the proliferation of a pest	1081	CAGACTACCGCTCTACTTATAGGCGATAAATGTGAAGCAACATGCAACACTCGTCG	1140
CC	CC	insect and potentially as medicinal delivery agents for cancer treatment	1081	CAGACTACCGCTCTACTTATAGGCGATAAATGTGAAGCAACATGCAACACTCGTCG	1140
CC	CC	and gene therapy.	1141	ACCAAGGCTTTTGGATCGCGCGCAGTACGCTCTCAACCCGACTAGCTACCGAGTACG	1200
XX	Sequence	2478 BP; 516 A; 854 C; 607 G; 501 T; 0 other;	1141	ACCAAGGCTTTTGGATCGCGCGCAGTACGCTCTCAACCCGACTAGCTACCGAGTACG	1200
Query Match	100.0%;	Score 2478; DB 19; Length 2478;	1201	ACGTGACGAGGCTACGCTCTGACACATTTTGACCTTCCGACAGACCATCGAGCGCGCTG	1260
Best Local Similarity	100.0%;	Pred. No. 0;	1201	ACGTGACGAGGCTACGCTCTGACACATTTTGACCTTCCGACAGACCATCGAGCGCGCTG	1260
Matches 2478;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;	1261	CAGTCCGCTTTGTGTGGGAGGTTTGCACAGAGGTGGCACTGCGCTCGAGGCACTCCAG	1320
QY	1	GTGTTTCTTTTACCAAGTGTGTAATAATTAACAAGAAAGAACACGACCTGTA	1261	CAGTCCGCTTTGTGTGGGAGGTTTGCACAGAGGTGGCACTGCGCTCGAGGCACTCCAG	1320
DB	1	GTGTTTCTTTTACCAAGTGTGTAATAATTAACAAGAAAGAACACGACCTGTA	1261	CAGTCCGCTTTGTGTGGGAGGTTTGCACAGAGGTGGCACTGCGCTCGAGGCACTCCAG	1320
QY	61	CCGCGCCCTTACACACTCGAGTCCGTCGACCGGATTTATACGTCGCCACACACGCG	1321	CCTGGAGCAGGATCTCTCGGTGGCTACCTCACTGCGCGCACAAACGGTACTATTTTCC	1380
DB	61	CCGCGCCCTTACACACTCGAGTCCGTCGACCGGATTTATACGTCGCCACACACGCG	1321	CCTGGAGCAGGATCTCTCGGTGGCTACCTCACTGCGCGCACAAACGGTACTATTTTCC	1380
QY	121	GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTTCGTCGACCAACCGGTTGGCA	1381	CAGCTGGCTCCGTTAGTACGTTCTCCCTGAGGGTTTCCGCTTGGAGCGCTACGACCGGA	1440
DB	121	GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTTCGTCGACCAACCGGTTGGCA	1381	CAGCTGGCTCCGTTAGTACGTTCTCCCTGAGGGTTTCCGCTTGGAGCGCTACGACCGGA	1440
QY	181	GTGACAGAGCTTCGGACCTAGAACCTCTCGAGGAGCGCACACACAGCACACA	1441	ACGACGGCTCTTTGGGACCGACTTCGCTTCCGAGGAGACACCGTCACTTTCCGGGAGGTG	1500
DB	181	GTGACAGAGCTTCGGACCTAGAACCTCTCGAGGAGCGCACACACAGCACACA	1441	ACGACGGCTCTTTGGGACCGACTTCGCTTCCGAGGAGACACCGTCACTTTCCGGGAGGTG	1500
QY	241	CCGCTTTAGCTGACCTACGCGAGCTTGTATAGCGGATTTATGAGCGGACACCAATC	1501	CGCTCGACAGAGTCTGTTGTGACAAACCCCGCGCGGCGGACGCCCCACCTTCA	1560
DB	241	CCGCTTTAGCTGACCTACGCGAGCTTGTATAGCGGATTTATGAGCGGACACCAATC	1501	CGCTCGACAGAGTCTGTTGTGACAAACCCCGCGCGGCGGACGCCCCACCTTCA	1560
QY	301	GCCCACTCCATCATTACCAACCCCGTTACCCCTTGCCCTTAATACCCCTGAACCTGAA	1561	CGGTGAGAGTGCCCTTCAAAACGGCTTACCAACACCGTGTGTTAGGAAACAGCTCTTAG	1620
DB	301	GCCCACTCCATCATTACCAACCCCGTTACCCCTTGCCCTTAATACCCCTGAACCTGAA	1561	CGGTGAGAGTGCCCTTCAAAACGGCTTACCAACACCGTGTGTTAGGAAACAGCTCTTAG	1620
QY	361	GCAGATGGAGATGCTGAGTGGCTGACAGGACCTCAACCGTTCGCGAACCGTA	1621	AGACTCGACCTCTCTCGTAGGCTCGAACTCCCTATGCCACCTGCTGACTTTGGACAGA	1680
DB	361	GCAGATGGAGATGCTGAGTGGCTGACAGGACCTCAACCGTTCGCGAACCGTA	1621	AGACTCGACCTCTCTCGTAGGCTCGAACTCCCTATGCCACCTGCTGACTTTGGACAGA	1680
QY	421	ACGTTCCGGTTCAGCGCAACACCGTCAACCGTCAATGGTAGAAGAACCAACCGGCTCGGA	1681	CGGTCCCAACAAACCCGAGATCGAGAGTCTGCTTTCTTAAAGAAACACTTGGCTCTATT	1740
DB	421	ACGTTCCGGTTCAGCGCAACACCGTCAACCGTCAATGGTAGAAGAACCAACCGGCTCGGA	1681	CGGTCCCAACAAACCCGAGATCGAGAGTCTGCTTTCTTAAAGAAACACTTGGCTCTATT	1740
QY	481	CCGGAAGGCAAGTTTCTCCCTGACAAATTTACCGCTGCTGCAACAAGACCTCGCGCAA	1741	TGCTTCACCTCCAAATGCGAAACCCCGTTTTCAGCTCAGCGCAGGAGCTCTTTGGG	1800
DB	481	CCGGAAGGCAAGTTTCTCCCTGACAAATTTACCGCTGCTGCAACAAGACCTCGCGCAA			
QY	541	GCCTTGACGCAACACCGTCACTTTCCCGCTTAAACATCTAGATGCGCGAATTCGGA			
DB	541	GCCTTGACGCAACACCGTCACTTTCCCGCTTAAACATCTAGATGCGCGAATTCGGA			
QY	601	ATTGGGCGAAGGAAGATCGACTCGATCCGATTTCCATCGGCTGGTACTTCAAGTACC			
DB	601	ATTGGGCGAAGGAAGATCGACTCGATCCGATTTCCATCGGCTGGTACTTCAAGTACC			
QY	661	TTGACCCAGCGGTGCTACAGAGTCTGCGCGCGCGTTCGCGAGTACTCGAAGATCCCTG			</

```
Db 1741 TGGTCCACTCAAAATGCGAAACCCCGTTTCCAGCTCACGCCAGCCAGCTCCTTTGGCG 1800
Qy 1801 CCGTTTCTCTCAACAATCCGGGTTATGAGCGCACACCGGACCTCCCGGACTACACTGGCA 1860
Db 1801 CCGTTTCTCTCAACAATCCGGGTTATGAGCGCACACCGGACCTCCCGGACTACACTGGCA 1860
Qy 1861 TCCGTGACTATTCGACAGAACATGTCCACCGGTGTGGGCCACTTCCTCGCTCACTCTCC 1920
Db 1861 TCCGTGACTATTCGACAGAACATGTCCACCGGTGTGGGCCACTTCCTCGCTCACTCTCC 1920
Qy 1921 ACTCTCGAGTATCGTCACTAAGACCTACCAAGGTTGGGAGGCGGTACGAACTCAACA 1980
Db 1921 ACTCTCGAGTATCGTCACTAAGACCTACCAAGGTTGGGAGGCGGTACGAACTCAACA 1980
Qy 1981 CGCTTTTCGGGCAATTCGCGCACCGGGCCCTCTCAAGATCAGGAGATCCTCTGCTCTCG 2040
Db 1981 CGCTTTTCGGGCAATTCGCGCACCGGGCCCTCTCAAGATCAGGAGATCCTCTGCTCTCG 2040
Qy 2041 CCGACGACTGGCCACCGCTCTCAGAGTGTCTACCCCGCCACTGACAACTTCGGGGCG 2100
Db 2041 CCGACGACTGGCCACCGCTCTCAGAGTGTCTACCCCGCCACTGACAACTTCGGGGCG 2100
Qy 2101 CCGTTTCTGCTTCCGCGGACATGCTGCTCCTCGTCTGAGTCCGAGGCAAGCTCT 2160
Db 2101 CCGTTTCTGCTTCCGCGGACATGCTGCTCCTCGTCTGAGTCCGAGGCAAGCTCT 2160
Qy 2161 CCATCATCAAGTCCGTTGGGAGACTGCGGCGCGGCTCAGTCCGGGCTCGCGAAGC 2220
Db 2161 CCATCATCAAGTCCGTTGGGAGACTGCGGCGCGGCTCAGTCCGGGCTCGCGAAGC 2220
Qy 2221 TACCGGACTCTAATGAGTGTACAGGAAAGATTGCCGCGGTGTCCGCGCGCGCGAG 2280
Db 2221 TACCGGACTCTAATGAGTGTACAGGAAAGATTGCCGCGGTGTCCGCGCGCGCGAG 2280
Qy 2281 CGCGCGCGCGCGCGCTGCTGCCAATAGTTGCTGCTCTCTGTTGCGCGCTTCGTAA 2340
Db 2281 CGCGCGCGCGCGCGCTGCTGCCAATAGTTGCTGCTCTCTGTTGCGCGCTTCGTAA 2340
Qy 2341 ACGGGCTGTCCGCGACATACGCTACCTAAAGACTCTGCTGAGTCCCGCTCGTTACA 2400
Db 2341 ACGGGCTGTCCGCGACATACGCTACCTAAAGACTCTGCTGAGTCCCGCTCGTTACA 2400
Qy 2401 CGACGGGTCTCCGCGGTTCGATTCCATTCCCAAGCGGCAAGAGGACGTAGTCTCT 2460
Db 2401 CGACGGGTCTCCGCGGTTCGATTCCATTCCCAAGCGGCAAGAGGACGTAGTCTCT 2460
Qy 2461 CGTCCCTCGGATACCA 2478
Db 2461 CGTCCCTCGGATACCA 2478
```

RESULT 2

AAQ58523

ID AAQ58523 standard; DNA; 2478 BP.

XX AC

XX AC

XX DT 25-MAR-2003 (updated)

XX DT 12-SEP-1994 (first entry)

DE Sequence of Heliothis armigera RNA 2 which encodes P17 and the
DE capsid proteins precursor P71.

XX HaSV; RNA 1; small RNA virus; P17; P71; ss.

XX Heliothis armigera stunt virus.

OS OS

XX FH

XX CDS

Key Location/Qualifiers

283..753

/*tag= a

/product= p17

366..2309

/*tag= b

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

/product= P71

W09404660-A1.

03-MAR-1994.

13-AUG-1993; 93WO-AU00411.

14-AUG-1992; 92AU-0004081.

08-JUL-1993; 93US-0089372.

(CSIR) COMMONWEALTH SCI & IND RES ORG.

(PACI-) PACIFIC SEEDS PTY LTD.

Christian PD, Gordon KHJ, Hanzlik TN;

WPI; 1994-083180/10.

P-PSDB; AAR49661, AAR49662.

Small RNA virus capable of infecting insect species, e.g.

Heliothis - and transgenic plants contg. viral nucleic acid, for

protection against insect pests

Disclosure; Figure 2; 183pp; English.

The inventors claim a virus comprising a genome hybridisable with

the nucleotide sequence of RNA 1 or RNA 2; pref. the sequences are

those given in Figs 1 and 2 of the specification. As isolated

protein or polypeptide prepn. of the proteins or polypeptides

derivable from the virus are also claimed.

H. armigera larvae were raised and viral RNA was extracted. The virus

RNAs were reverse transcribed into cDNA. Clone hr236 contains about

88% or RNA 2. RNA 2 encodes a protein of mol.wt.71,000(P71) which

contains the peptide sequences corresp. to those determined from

the two virus capsid proteins. This protein is therefore the

precursor of these capsid proteins. In addition, another major

translation product of apparent mol. wt. 24,000 is obtd. This

protein is derived from a mol. wt. 17,000 reading frame overlapping

the slab of the capsid protein gene. The Mr 24,000 protein (referred

to as P17) may have a function in modifying or manipulating the growth

characteristics or cell cycle of HaSV-infected cells.

(Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 2478 BP; 516 A; 853 C; 608 G; 501 T; 0 other;

Query Match 99.6%; Score 2468.4; DB 15; Length 2478;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2472; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 GTTTTCTTTCTTTTACCAAGTGTGTAATAATTTAAACAAAGAAAGAAACCGGTA 60

1 GTTTTCTTTCTTTTACCAAGTGTGTAATAATTTAAACAAAGAAAGAAACCGGTA 60

61 CCGGCGCTTTACACACTCGAGTCCGTCGACACCGGATTATACGTCGCCACACACGCG 120

61 CCGGCGCTTTACACACTCGAGTCCGTCGACACCGGATTATACGTCGCCACACACGCG 120

121 GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTTCTGTCGACACCGGTTGGCA 180

121 GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTTCTGTCGACACCGGTTGGCA 180

181 GTCCACAGAGCGTTCCCGGACCACTAGAACCTCTCGAGCGACGACACACACACACA 240

181 GTCCACAGAGCGTTCCCGGACCACTAGAACCTCTCGAGCGACGACACACACACA 240

241 CCGCTTTAGCTGCACTACCGGAGTGTAGTAGCGCGGATTTATGAGCGAGCACCATC 300

241 CCGCTTTAGCTGCACTACCGGAGTGTAGTAGCGCGGATTTATGAGCGAGCACCATC 300

301 GCCCAGCTCCATCATTTACCAACCGGTTACACCCCTTGCCTTAATACCCCTGAACCTGAA 360

301 GCCCAGCTCCATCATTTACCAACCGGTTACACCCCTTGCCTTAATACCCCTGAACCTGAA 360

QY 361 CGAGATGGAGATGCTGGAGTGGCTCACAGCGACCTCACAAACCGTCGCGGAACCCGTA 420
DB 361 CGAGATGGAGATGCTGGAGTGGCTCACAGCGACCTCACAAACCGTCGCGGAACCCGTA 420
QY 421 ACGTTGCGGTGAGGCGCAACACCGTACCGTCAATGGTGAAGAAAACCAACGGGCTGGA 480
DB 421 ACGTTGCGGTGAGGCGCAACACCGTACCGTCAATGGTGAAGAAAACCAACGGGCTGGA 480
QY 481 CCGGAAGCAAGTTTCTCCCTTGCACATTTACCGGTGCTGCAACAGACCTCGCGAAA 540
DB 481 CCGGAAGCAAGTTTCTCCCTTGCACATTTACCGGTGCTGCAACAGACCTCGCGAAA 540
QY 541 GCCTTGAGCCCAACACCGTCACTTTCCCGCTAACATCTTAGCATGCTGCGGAAA 600
DB 541 GCCTTGAGCCCAACACCGTCACTTTCCCGCTAACATCTTAGCATGCTGCGGAAA 600
QY 601 ATTGGGCAAGGGAAGATCGACTCGACTCGGATTCCTGATTCCTGATTCCTGATTC 660
DB 601 ATTGGGCAAGGGAAGATCGACTCGACTCGGATTCCTGATTCCTGATTCCTGATTC 660
QY 661 TTGACCCAGCGGTGCTACAGATCTGCGCGCGCTCGGCGAGTACTCGAAGATCCCTG 720
DB 661 TTGACCCAGCGGTGCTACAGATCTGCGCGCGCTCGGCGAGTACTCGAAGATCCCTG 720
QY 721 ACGGCTCGTCAAGTTCTCCGTCGACGAGATAAGAGATCTATAACGAGGAGTCC 780
DB 721 ACGGCTCGTCAAGTTCTCCGTCGACGAGATAAGAGATCTATAACGAGGAGTCC 780
QY 781 CCGTCTGCTCACTGAGTGTCCCTCGACGCGCGCTCGGCGAGTACTCGAAGATCCCTG 840
DB 781 CCGTCTGCTCACTGAGTGTCCCTCGACGCGCGCTCGGCGAGTACTCGAAGATCCCTG 840
QY 841 CTTTTCGATGTTGAGAACCGCTACGTCGCGCTAGCGAAGCTCGAGAACGAGATGT 900
DB 841 CTTTTCGATGTTGAGAACCGCTACGTCGCGCTAGCGAAGCTCGAGAACGAGATGT 900
QY 901 CGCTCGAGTTGTCAACGACCTCATCGAGTGGCTCAACAACTCGCGGCTGGGTTATG 960
DB 901 CGCTCGAGTTGTCAACGACCTCATCGAGTGGCTCAACAACTCGCGGCTGGGTTATG 960
QY 961 TCGTTGACTTGAACAGTGGATTAACTTCAACAATGACACCACTGATCTCGCGATCC 1020
DB 961 TCGTTGACTTGAACAGTGGATTAACTTCAACAATGACACCACTGATCTCGCGATCC 1020
QY 1021 CGGTTCTAGTCCAACTACAGCTTCAGACCTTCAGACCTTCAGAGGCTGTTGCGACAGTCT 1080
DB 1021 CGGTTCTAGTCCAACTACAGCTTCAGACCTTCAGACCTTCAGAGGCTGTTGCGACAGTCT 1080
QY 1081 CAGACTACCGCTCACTTATAAGGCGATAACATGTGAAGCCAAACATGCAACACTCGTCTG 1140
DB 1081 CAGACTACCGCTCACTTATAAGGCGATAACATGTGAAGCCAAACATGCAACACTCGTCTG 1140
QY 1141 ACCAAGGCTTTGGATCGGCGCGAGTACGCTCTCACCCGACTAGCTACCGCAGTACG 1200
DB 1141 ACCAAGGCTTTGGATCGGCGCGAGTACGCTCTCACCCGACTAGCTACCGCAGTACG 1200
QY 1201 ACGTCAGGAGGCTACGCTCTGACACTTTGACCTTCGCGAGACCTACGAGCGCGCTG 1260
DB 1201 ACGTCAGGAGGCTACGCTCTGACACTTTGACCTTCGCGAGACCTACGAGCGCGCTG 1260
QY 1261 CACTCGGCTTTGTGGGACGTTTGGCCAGGTTGGCCAGGTTGGCCTTCGAGGCACTCCAG 1320
DB 1261 CACTCGGCTTTGTGGGACGTTTGGCCAGGTTGGCCAGGTTGGCCTTCGAGGCACTCCAG 1320
QY 1321 CTTGGGAGCAGGCACTCTCGGCTGCTACCTCACTGCGCGCAACAAAGGTACTACTTTCC 1380
DB 1321 CTTGGGAGCAGGCACTCTCGGCTGCTACCTCACTGCGCGCAACAAAGGTACTACTTTCC 1380
QY 1381 CAGTGGCTCGTTAGCTACGTTCTCCCTGAGGTTTCGCCCTTGGAGGCTACGACCCGA 1440
DB 1381 CAGTGGCTCGTTAGCTACGTTCTCCCTGAGGTTTCGCCCTTGGAGGCTACGACCCGA 1440
QY 1441 ACGACGGCTCTTGGACCGACTTGGCTTCGCGAGGAGACACCGTCACTTTCGCGAGGTCG 1500

DB 1441 ACGACGGCTCTTGGACCGACTTGGCTTCGCGAGGAGACACCGTCACTTTCGCGAGGTCG 1500
QY 1501 CCGTCGACGAGGTCGTTGTGACCAACACCCCGCGCGCGGAGCGGCGCCACCTTCA 1560
DB 1501 CCGTCGACGAGGTCGTTGTGACCAACACCCCGCGCGCGGAGCGGCGCCACCTTCA 1560
QY 1561 CCGTGAGAGTGCCCGCTTCAACCGCTTACCAACACCGGTTTGAAGAACCGCTCTTAG 1620
DB 1561 CCGTGAGAGTGCCCGCTTCAACCGCTTACCAACACCGGTTTGAAGAACCGCTCTTAG 1620
QY 1621 AGACTCGACCGCTCTCTGCTAGGTCGAATCCTGATGCGACCTGCTGACTTTGGACAGA 1680
DB 1621 AGACTCGACCGCTCTCTGCTAGGTCGAATCCTGATGCGACCTGCTGACTTTGGACAGA 1680
QY 1681 CCGTGCGCAACACCCGAGATCGAGCAGTCTTTAAAGAAAACACTTTGGCTGCTATT 1740
DB 1681 CCGTGCGCAACACCCGAGATCGAGCAGTCTTTAAAGAAAACACTTTGGCTGCTATT 1740
QY 1741 TGGTCCACTCCAAATTCGGAACCCCGTTTTCCAGCTCACGCCAGCAGCTCTTTGGCG 1800
DB 1741 TGGTCCACTCCAAATTCGGAACCCCGTTTTCCAGCTCACGCCAGCAGCTCTTTGGCG 1800
QY 1801 CCGTTTCTTCAACATTCGGGTTATGAGCGCAACCGGACCTCCCGGACTACACTGGCA 1860
DB 1801 CCGTTTCTTCAACATTCGGGTTATGAGCGCAACCGGACCTCCCGGACTACACTGGCA 1860
QY 1861 TCGTGACTCATTCGACCAACATGTCACCGCTGCGGCGGCTTCCGCTCACTCTCCC 1920
DB 1861 TCGTGACTCATTCGACCAACATGTCACCGCTGCGGCGGCTTCCGCTCACTCTCCC 1920
QY 1921 ACTCTGACTATCTGCTCACTAAGACCTTACAGGCTTGGGAAGGCTCACGAAACGTCAACA 1980
DB 1921 ACTCTGACTATCTGCTCACTAAGACCTTACAGGCTTGGGAAGGCTCACGAAACGTCAACA 1980
QY 1981 CGCTTTTCGCGCAATTCGCGACCGGCTCTCTCAAGAAATGAGAGATCTCTGCGCTCG 2040
DB 1981 CGCTTTTCGCGCAATTCGCGACCGGCTCTCTCAAGAAATGAGAGATCTCTGCGCTCG 2040
QY 2041 CCGACGACTGGGACCCGCTCTACAGGCTGTACCCCGGCTGACAACTTCGCGGCGG 2100
DB 2041 CCGACGACTGGGACCCGCTCTACAGGCTGTACCCCGGCTGACAACTTCGCGGCGG 2100
QY 2101 CCGTTTCTGCTTCGCGCGAAACATGCTCTCTCGCTGCTGAACTCGGAGCAACGTCT 2160
DB 2101 CCGTTTCTGCTTCGCGCGAAACATGCTCTCTCGCTGCTGAACTCGGAGCAACGTCT 2160
QY 2161 CCATCATCAAGTCGTTTGGGAGACTGCGCTCGCGCGGCTCAGTTCGCGGCTCGCGAAGC 2220
DB 2161 CCATCATCAAGTCGTTTGGGAGACTGCGCTCGCGCGGCTCAGTTCGCGGCTCGCGAAGC 2220
QY 2221 TACCGGACTGCTAATGAGTGTACAGGGAAGATTGCGCGGCTGTCGCGCGGCGGAG 2280
DB 2221 TACCGGACTGCTAATGAGTGTACAGGGAAGATTGCGCGGCTGTCGCGCGGCGGAG 2280
QY 2281 CCGCGCGCGCGCTGCGCAATTAGTTTGTCTCTCTCTGTTTCGCGGTTTCGTAATA 2340
DB 2281 CCGCGCGCGCGCTGCGCAATTAGTTTGTCTCTCTCTGTTTCGCGGTTTCGTAATA 2340
QY 2341 ACGCGTGGTCCGCGCACTTACCGGTAACCTTAAAGACTCTGGTGAGTCCCGCTGTTACA 2400
DB 2341 ACGCGTGGTCCGCGCACTTACCGGTAACCTTAAAGACTCTGGTGAGTCCCGCTGTTACA 2400
QY 2401 CGACGGTCTGCGCGGTTGATTCATTCGAGCGGCAAGGAGGAGTGTAGTCTCT 2460
DB 2401 CGACGGTCTGCGCGGTTGATTCATTCGAGCGGCAAGGAGGAGTGTAGTCTCT 2460
QY 2461 GCGTCCCTCGGGAATACCA 2478
DB 2461 GCGTCCCTCGGGAATACCA 2478

RESULT 3

XX PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA -
XX
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 985 BP; 234 A; 145 C; 326 G; 280 T; 0 other;
Query Match 2.0%; Score 49; DB 24; Length 985;
Best Local Similarity 47.5%; Pred. No. 0.035;
Matches 145; Conservative 0; Mismatches 160; Indels 0; Gaps 0;
QY 720 GACGCTCGTCAAGTTCTCGTCGACGAGAGATAGAGATCTATACGAGGATGC 779
DB 715 GACGTCACGTCACATCTACATCGACATCAACATCTAGTCGACATCTACGTCGACATC 656
QY 780 CCGCTCGTCACGTCGAGTTCGCTCCGTCGACGCGCGGCGGAGCTCTCGATTTTC 839
DB 655 GAGTCGACATCGACATCCGCTCTACATCTAGCTCAAGTCTACATCTACGTTACATC 596
QY 840 TCGTTTCCGATGTTTCAGAACCGCTTACGTCGCGGTCGAGAACGTCGAGAACGAGATG 899
DB 595 GACGTCGACATCTAGCTCAACGTCACGTCGACATCGACGTCGACATCTACGTCGACGTC 536
QY 900 TCGTCGAGTGTCAACGACCTCATCGAGTGGCTCAACATCTCCCGACTGGCGTTAT 959
DB 535 TACATCTAGCTCGAGTCGACGTCATCATCTAGCTTACATCAACATCAACATCTACATC 476
QY 960 GTCGTTGACTCTGAACAGTGGATTAACTTCAACATGACACACGTCATCTACGTCGACATC 1019
DB 475 TACATCTAGCTCGAGTCTACATCTAGCTCGACATCTAGCTTACATCTACGTTACATC 416
QY 1020 CCGGT 1024
DB 415 TACGT 411

RESULT 7
ABQ44817
ID ABQ44817 standard; DNA; 985 BP.
XX ABQ44817;
XX
XX 12-JUL-2002 (first entry)
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 31408.
XX
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.

XX OS Homo sapiens.
XX PN WO200218632-A2.
XX PD 07-MAR-2002.
XX PF 01-SEP-2001; 2001WO-EP10074.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PR 05-SEP-2000; 2000DE-1044543.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.
XX PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA -
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 985 BP; 280 A; 326 C; 145 G; 234 T; 0 other;
Query Match 2.0%; Score 49; DB 24; Length 985;
Best Local Similarity 47.5%; Pred. No. 0.035;
Matches 145; Conservative 0; Mismatches 160; Indels 0; Gaps 0;
QY 720 GACGCTCGTCAAGTTCTCGTCGACGAGAGATAGAGATCTATACGAGGATGC 779
DB 271 GACGTCACGTCACATCTACATCGACATCAACATCTAGTCGACATCTACGTCGACATC 330
QY 780 CCGCTCGTCAAGTTCGCTCCGTCGACGCGCGGCGGAGCTCTCGATTTTC 839
DB 331 GAGTCGACATCGACATCCGCTCTACATCTAGCTCAAGTCTACATCTACGTTACATC 390
QY 840 TCGTTTCCGATGTTTCAGAACCGCTTACGTCGCGGTCGAGAACGTCGAGAACGAGATG 899
DB 391 GACGTCGACATCTAGCTCAACGTCACGTCGACATCGACGTCGACATCTACGTCGACGTC 450
QY 900 TCGTCGAGTGTCAACGACCTCATCGAGTGGCTCAACATCTCCCGACTGGCGTTAT 959
DB 451 TACATCTAGCTCGAGTCGACGTCATCATCTAGCTTACATCAACATCAACATCTACATC 510
QY 960 GTCGTTGACTCTGAACAGTGGATTAACTTCAACATGACACACGTCATCTACGTCGACATC 1019
DB 511 TACATCTAGCTCGAGTCTACATCTAGCTCGACATCTAGCTTACATCTACGTTACATC 570
QY 1020 CCGGT 1024
DB 415 TACGT 411

```
Db 571 TACGT 575
RESULT 8
ID AAI99683
XX AAI99683 standard; DNA; 4403765 BP.
AC AAI99683;
XX
XX 15-JAN-2002 (first entry)
XX
XX Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
XX Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
XX variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
XX Mycobacterium tuberculosis.
XX
XX US6294328-B1.
XX
XX 25-SEP-2001.
XX
XX 24-JUN-1998; 98US-0103840.
XX
XX 24-JUN-1998; 98US-0103840.
XX
XX (GENO-) INST GENOMIC RES.
XX
XX Fleischmann RD, White OR, Fraser CM, Venter JC;
XX WPI; 2001-647261/74.
XX
XX Evaluating strain variation of Mycobacterium tuberculosis, comprises
XX determining the nucleotide sequence of the strain at positions in the
XX genome corresponding to positions where M. tuberculosis strains CDC
XX 1551 and H37Rv differ -
XX
XX Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
XX The invention relates to evaluating strain variation within and between
XX different populations of the tuberculosis bacterial pathogen,
XX Mycobacterium tuberculosis or related Mycobacterium by determining the
XX nucleotide sequence of the first strain at positions in the complete
XX sequence of the genome that correspond to positions that differ in the
XX nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
XX H37Rv (AAI99682). The method is useful for evaluating strain variation of
XX M. tuberculosis and has valuable application in the fields of
XX tuberculosis genetics, epidemiology, patient treatment and epidemic
XX monitoring.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
XX Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
XX
XX Query Match 2.0%; Score 49; DB 22; Length 4403765;
XX Best Local Similarity 47.3%; Pred. No. 0.65;
XX Matches 148; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
XX
XX 1251 AGCGCGCTGCACTCGCTTTGTGGCAGAGGTTTGCACAGAGTGGCCTGCA 1310
XX 1213114 ATCGCGCGGAGACCCGCGCGCGCGCGGTGGGGCGGAGGCAACGGTGGGCGGA 1213173
XX
XX 1311 GGCACCTCCAGCTGGAGCAGGATCTCGGGTGGTACTCCTCAGCTGGCGCCACACGGT 1370
XX 1213174 GGCACGGTGGGTGGTGTGGCAATGGTGGTGGCGCGGTTCGGTGGCGCGCGC 1213233
XX
XX 1371 ACTACTTCCAGCTCGCTCGGTAGCTAGCTTCTCCCTGAGGGTTTCGCCCTTGAGCG 1430
XX 1213234 GCGCGCGCGGTGGCGGTGCCCGCAACTTGGGCTCGCGCGGGGGTATCAACGCCCGCGC 1213293
XX
XX 1431 TAGACCCGCAACGACGGCTCTTGGACCGACTTCGCTTCGCGAGGAGACACCGTCACTTTC 1490
XX
Db 1213294 GGTAAACCCCGAGCGGCTCGTGGCATCGCGGTGTCGGCGGCGCACCGCC 1213353
QY 1491 CGGCAGGTGCGCGTCTGAGAGGTCTGTTGTACCAACAACCCCGCGGCGGCGAGCGCC 1550
Db 1213354 GGGCTGTTTCGGCGAGCGTGGGGCTGGTGGCGCGCGCGCGCGGCTTC 1213413
QY 1551 CCCACCTTTCACCG 1563
Db 1213414 GCGGCATCAGG 1213426
RESULT 9
AAI99682
ID AAI99682 standard; DNA; 4411529 BP.
XX
XX AC AAI99682;
XX
XX 15-JAN-2002 (first entry)
XX
XX Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
XX Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
XX variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
XX Mycobacterium tuberculosis.
XX
XX US6294328-B1.
XX
XX 25-SEP-2001.
XX
XX 24-JUN-1998; 98US-0103840.
XX
XX 24-JUN-1998; 98US-0103840.
XX
XX (GENO-) INST GENOMIC RES.
XX
XX Fleischmann RD, White OR, Fraser CM, Venter JC;
XX WPI; 2001-647261/74.
XX
XX Evaluating strain variation of Mycobacterium tuberculosis, comprises
XX determining the nucleotide sequence of the strain at positions in the
XX genome corresponding to positions where M. tuberculosis strains CDC
XX 1551 and H37Rv differ -
XX
XX Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
XX The invention relates to evaluating strain variation within and between
XX different populations of the tuberculosis bacterial pathogen,
XX Mycobacterium tuberculosis or related Mycobacterium by determining the
XX nucleotide sequence of the first strain at positions in the complete
XX sequence of the genome that correspond to positions that differ in the
XX nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
XX H37Rv (AAI99682). The method is useful for evaluating strain variation of
XX M. tuberculosis and has valuable application in the fields of
XX tuberculosis genetics, epidemiology, patient treatment and epidemic
XX monitoring.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
XX Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;
XX
XX Query Match 2.0%; Score 49; DB 22; Length 4411529;
XX Best Local Similarity 47.3%; Pred. No. 0.65;
XX Matches 148; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
XX
XX 1251 AGCGCGCTGCACTCGCTTTGTGGCAGAGTGGTGGCAGAGTGGCCTGCA 1310
XX 1212977 ATCGCGCGGAGACCCGCGCGCGCGGTGGGGCGGAGGCAACGGTGGGCGCGGA 1213036
XX
XX 1311 GGCACCTCCAGCTGGAGCAGGATCTCGGGTGGTACTCCTCAGCTGGCGCCACACGGT 1370
XX
QY 1311 GGCACCTCCAGCTGGAGCAGGATCTCGGGTGGTACTCCTCAGCTGGCGCCACACGGT 1370
```

Db 1213037 GGCAAGGTGGGTGTTGTTGGCAATGTTGGCGGGGCTCCGGTGGCGACGGCGGC 1213096
 QY 1371 ACTACTTTCCAGCTGGCTCCGTTAGCTACTGTTCTCCCTGAGGGTTTCGCCCTTGAGGCG 1430
 Db 1213097 GCGCGCGGCGTGGCGTCCGCAACTTGGGCTCGCGCGGGGTATCAACGCCCGCGCC 1213156
 QY 1431 TAGGACCCGAAGCAGCGCTCTTGGACCGACTTCGCTTCGGCAGGAGACACCGTCACTTTC 1490
 Db 1213157 GGTAACCCCGGAGCGGCTCGGTTCGGCATCGCGGGTGGCGGTGGTCCGCGGCAACCGCC 1213216
 QY 1491 CGCGAGGTGGCTGACGAGGTGCTGTGACCAACACCCCGCGGGCGGCGGCGCGCC 1550
 Db 1213217 GGGCTGTTCCGCGACGCTGGGCTGCTGGGCGCGGTGGTGGCGCGCGCGGCTTC 1213276
 QY 1551 CCCACCTTCACCG 1563
 Db 1213277 GCGCGCATCAGCG 1213289

RESULT 10

AAAL0594/c

ID AAAL0594 standard; DNA; 10732 BP.

XX

AC AAAL0594;

XX 29-JUN-2000 (first entry)

XX DE Gene encoding a subunit of cellulose synthase.

XX KW Cellulose synthase; cellulose production; increase yield; ds.

XX OS Vigna angularis.

XX PN JP2000060568-A.

XX PD 29-FEB-2000..

XX PF 26-AUG-1998; 98JP-0239998.

XX PR 26-AUG-1998; 98JP-0239998.

XX (MIZU/) MIZUNO K.

XX PA (OJIP) OJI PAPER CO.

XX DR WPI; 2000-342371/30.

XX P-PSDB; AA85179.

XX A gene encoding a cellulose synthetic equipment - for the improvement
 in the amount of cellulose synthesised in a plant body
 Claim 2; Page 14-21; 32pp; Japanese.
 This sequence represents a gene encoding a subunit of the cellulose
 synthase complex of Vigna angularis. The invention relates to subunits of
 cellulose synthetic equipment, that can be used to increase the amount of
 cellulose synthesised by a plant. The proteins and genes encoding them
 can also be used to improve the properties of the cellulose being
 produced by a plant.

XX SQ Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 2251 other;

Query Match 1.9%; Score 47.6; DB 21; Length 10732;

Best Local Similarity 14.6%; Pred. No. 0.18;

Matches 115; Conservative 312; Mismatches 351; Indels 7; Gaps 3;

QY 1386 GGCCTCGTTAGTACGTTCTCCCTGAGGGTTTCGCCCTTGAGCGCTAGCACCGCACGAC 1445
 Db 10057 SRGTBDDCCTBSRSCYSYDASRYANCDTBCYTTBAKYRACYDAYAKRONSTSRAY 9998
 QY 1446 GGCCTCTGGACGACTTCGCTCCGCGAGGAGACACGTCACCTTCGGGACGCTCGCGTC 1505
 Db 9997 SDSTCYTRCRCRNCYSTYSYSTVRASTTBTTCYTCTBCSRRCYSRYSSTCNCYSYC 9938

QY 1506 GACGAGTCTGTGACCAACCAACCCCGCGCG--GCGGAGCGCCCGCCACCTTCACCG 1563
 Db 9937 CYTSRYSTTNNSTCRCTTYSYNTTBTBYSYSTYSRGSRSRSDSRGNCYVNSTNCY 9878
 QY 1564 TGAGAGTCCCGCTTCAACGCTTACACCAACCGCTGTTTAGGAACACGCTCTTAGAGA 1623
 Db 9877 DASTSDTBSRCYTYSYSTSDSTNSTTBSDCYTTTBSRSTSDSTSTFRCSS 9818
 QY 1624 CTGACCCCTCTCTCGTAGGCTCGAATCTCCCTATGCCACCTGCTGCTACTTTGGACAGCG 1683
 Db 9817 DYDATBSNSTNCCYDASRTBTBNTCYVBYDARCSRDSYSSRGYDANSYSS 9758
 QY 1684 TCGCAACCAACCCGAGATCGAGAGTCGCTTCTTAAAGAAACACTTGGCTGCTATTGG 1743
 Db 9757 RYS----STYSSDSYSTSAKYCAKSTTBTBYDAYDACYDAYDANCYSDSTYBYCSRR 9702
 QY 1744 TCCACTCCAAATGCAACCCCGTTTTCAGCTCACGCCAGCAGCTCTTTGGCGCG 1803
 Db 9701 CCCYDAYSCRYDARCYDACYSYNSNCTYDATTBTTSRYSTTYSYNCYDATT--SRCTB 9643
 QY 1804 TTTCCTTCAACAATCCGGGTTATGAGCGCACACGCGACTCCCGGACTACACTGGCATCC 1863
 Db 9642 YSTBTBTSRCAKCTBDSTSTAKNSTSTYTTCTBYSRSGYSCSRSRRCYSCYTD 9583
 QY 1864 GTGACTCATTCGACCAACATGTCACCGCTGTGGCCACTTCGCTCACTCTCCACT 1923
 Db 9582 SDSTCYSTTSTTAYSCYTSRGYSYDASRSTSTYSRCTTSTYSYSTTDSYSDYSTTT 9523
 QY 1924 CTGCACTATCGTCACTAAGACTCAGGCTTGGGAGGCGTCAAGACGTCACACCG 1983
 Db 9522 ENSTYSSDCTBYSSDRCSRSDSTCNCYSYSDSRSTYDACYTYDAKTBCYTYSDNC 9463
 QY 1984 CTTTCGGCAATTTCGCGCACCGGCTCTCAAGAATGAGGAGATCTCTGCTCCGCG 2043
 Db 9462 CNSTSRCTNSTYSCSRBTBSRSTCCCTBTTSRGNCCYDAYDANSTSYDAYDACYSY 9403
 QY 2044 ACGACTGGCACCGCTCTCAGGTGTCTACCGCCCACTGACAACTTCGCGCGCGCG 2103
 Db 9402 DASTBYSYCTBYSNSTYDAYSSRYSCYTCYDYSSTCTCAKCTBCNSTSRKST 9343
 QY 2104 TTTCTGCTTCGCGCGCAACATGCTCTCGTCTGAAGTCGAGCGCAAGCTCTCCA 2163
 Db 9342 TTTTTCYCTSYCTSYTTCYTCYTNCSRGCCSRSTTSRSDYDASCNCSRSDYSTBCT 9283
 QY 2164 TCATC 2168
 Db 9282 TYDAC 9278

RESULT 11

AAI99683/c

ID AAI99683 standard; DNA; 4403765 BP.

XX AC AAI99683;

XX DT 15-JAN-2002 (first entry)

XX DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.

XX KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
 variation; epidemiology; patient treatment; epidemic monitoring; ds.
 XX OS Mycobacterium tuberculosis.
 XX FN US6294328-B1.
 XX PD 25-SEP-2001.
 XX PF 24-JUN-1998; 98US-0103840.
 XX PR 24-JUN-1998; 98US-0103840.
 XX PA (GENO-) INST GENOMIC RES.

```
XX PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX PA WPI; 2001-647261/74.
XX DR
XX
XX PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
XX PT determining the nucleotide sequence of the strain at positions in the
XX PT genome corresponding to positions where M. tuberculosis strains CDC
XX PT 1551 and H37Rv differ
XX
XX PS Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
XX CC The invention relates to evaluating strain variation within and between
XX CC different populations of the tuberculosis bacterial pathogen,
XX CC Mycobacterium tuberculosis or related Mycobacterium by determining the
XX CC nucleotide sequence of the first strain at positions in the complete
XX CC sequence of the genome that correspond to positions that differ in the
XX CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
XX CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
XX CC M. tuberculosis and has valuable application in the fields of
XX CC tuberculosis genetics, epidemiology, patient treatment and epidemic
XX CC monitoring.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from USPTO
XX CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
Query Match 1.9%; Score 46; DB 22; Length 4403765;
Best Local Similarity 48.1%; Pred. No. 3.6;
Matches 130; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 1935 GTCACCTAAGACCTACACAGGTTGGGAAGCGTCACGAACGTCAACACGCTTTTCGGCCAA 1994
DB 2534569 GTCACACACGCGCTGTCGGACGCTTTCTCGTCCCGGTGTGAGACGCTGTCGCGCAA 2534510
QY 1995 TTCCGCGCAGCGGGCGCTCTCAAGAATGAGGAGATCTCTGCTCCGCGCAGCAGCTGGCC 2054
DB 2534509 TTGACGGCGCGCTTTCGGCGGCGGTGTGCGGCGCAGCAGACCCGCGCAGCAGCGCA 2534450
QY 2055 ACCCGTCTCACAGGTGTCTACCCCGCCTGACAACTTCGGCGCGCGCGCTTTCTGCTTTC 2114
DB 2534449 ACGGTGCTGACCGCGTCACTTCGGCGCAGCGCGCGCGCGCGATGCGCGCGCTCG 2534390
QY 2115 GCCCGCAACATGCTGCTCCGCTGAGTGTGGAGCAACGTCCTCATCAAGTCC 2174
DB 2534389 GTCTGACGCGCGCTTTGGCGTGTGCGAGGCGCAGGAGACTCCACATCGGCGCGCC 2534330
QY 2175 GTTGGCGAGACTGCGCTGGCGCGGCTCAG 2204
DB 2534329 CCTGGCTACACGCGCGCGCGCGCTGCTG 2534300
RESULT 12
AAI99682/c
ID AAI99682 standard; DNA; 4411529 BP.
XX
XX AC AAI99682;
XX
XX DT 15-JAN-2002 (first entry)
XX
XX DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
XX KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
XX KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
XX OS Mycobacterium tuberculosis.
XX
XX PN US6294328-B1.
XX
XX PD 25-SEP-2001.
XX
XX PF 24-JUN-1998; 98US-0103840.
```

```
XX 24-JUN-1998; 98US-0103840.
XX (GENO-) INST GENOMIC RES.
XX
XX PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX DR WPI; 2001-647261/74.
XX
XX PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
XX PT determining the nucleotide sequence of the strain at positions in the
XX PT genome corresponding to positions where M. tuberculosis strains CDC
XX PT 1551 and H37Rv differ
XX
XX PS Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
XX CC The invention relates to evaluating strain variation within and between
XX CC different populations of the tuberculosis bacterial pathogen,
XX CC Mycobacterium tuberculosis or related Mycobacterium by determining the
XX CC nucleotide sequence of the first strain at positions in the complete
XX CC sequence of the genome that correspond to positions that differ in the
XX CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
XX CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
XX CC M. tuberculosis and has valuable application in the fields of
XX CC tuberculosis genetics, epidemiology, patient treatment and epidemic
XX CC monitoring.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from USPTO
XX CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;
Query Match 1.9%; Score 46; DB 22; Length 4411529;
Best Local Similarity 48.1%; Pred. No. 3.6;
Matches 130; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 1935 GTCACCTAAGACCTACACAGGTTGGGAAGCGTCACGAACGTCAACACGCTTTTCGGCCAA 1994
DB 2537400 GTCACACGACGCTGTCGGACGCTTTCTCGTCCCGGTGTGAGACGCTGTCGCGCAA 2537341
QY 1995 TTCCGCGCAGCGGGCGCTCTCAAGAATGAGGAGATCTCTGCTCCGCGCAGCAGCTGGCC 2054
DB 2537340 TTGACGCGCGCTTTTCGGCGGCGGTGTGCGGCGCAGCAGACCCGCGCAGCAGCGCA 2537281
QY 2055 ACCCGTCTCACAGGTGTCTACCCCGCCTGACAACTTCGCGCGCGCGCTTTCTGCTTTC 2114
DB 2537280 ACGGTGCTGACCGCGCTGCTGCTGCTGAGTGTGGAGCAACGTCCTCATCAAGTCC 2174
QY 2115 GCCCGCAACATGCTGCTCCGCTGAGTGTGGAGCAACGTCCTCATCAAGTCC 2174
DB 2537220 GTCTTGCAGCGCGCTTTGGCGTGTGCGAGGCGCAGGAGACTCCACATCGGCGCGCC 2537161
QY 2175 GTTGGCGAGACTGCGCTGGCGCGGCTCAG 2204
DB 2537160 CCTGGCTACACGCGCGCGCGCGCTGCTG 2537131
RESULT 13
ACC44572
ID ACC44572 standard; cDNA; 3285 BP.
XX
XX AC ACC44572;
XX
XX DT 02-JUN-2003 (first entry)
XX
XX DE Alpha-amylase/glucoamylase fusion nucleotide sequence SEQ ID NO:46.
XX
XX KW Self-processing plant; plant; processing enzyme; alpha-amylase; grain;
XX KW pullulanase; alpha-glucosidase; glucose isomerase; glucoamylase;
XX KW mesophilic; thermophilic; hyperthermophilic; transgenic plant; starch;
XX KW maltodextrin; ethanol; fermentation; beverage; enzyme; gene; ss.
XX
XX OS Aspergillus shirousami.
```

OS Synthetic.
XX Key Location/Qualifiers
FH 1..3285
FT /*tag=a
FT /partial
FT /product= "alpha-amylase/glucoamylase fusion protein"
FT /note= "no start or stop codons given"
XX WO2003018766-A2.
XX
XX 06-MAR-2003.
XX
XX 27-AUG-2002; 2002WO-US27129.
XX
XX 27-AUG-2001; 2001US-315281P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Lanahan MB, Basu SS, Batie CJ, Chen W, Craig J, Kinkema M;
XX
XX WPI; 2003-268420/26.
XX P-PSDB; ABP96630.
XX
XX Novel polynucleotide encoding hyperthermophilic processing enzymes e.g.
PT alpha-amylase, useful for producing plant to produce food products
PT having improved taste or fermentable substrates for ethanol
XX
XX Claim 1; Page 107-109; 158pp; English.
XX
XX The present invention describes polynucleotides which encode processing
CC enzymes (e.g. alpha-amylase, pullulanase, alpha-glucosidase, glucose
CC isomerase, or glucoamylase) that are optimised for expression in plants.
CC The polynucleotides encode mesophilic, thermophilic or hyperthermophilic
CC processing enzymes, which are activated under suitable conditions to act
CC upon the desired substrate. Also described are self-processing transgenic
CC plants and plant parts, e.g. grain, which express one or more of these
CC enzymes and have an altered composition that facilitates plant and grain
CC processing. Also described is a method (M) for converting starch to
CC starch-derived products in a transformed plant part (pp), by activating
CC the starch processing enzyme contained in it. Transgenic grain is useful
CC for preparing maltodextrin. A transformed plant (TP) can be used to
CC produce food products having improved taste and to produce fermentable
CC substrates for ethanol and fermented beverages. (M) eliminates the need
CC to mill or physically disrupt the integrity of plant parts prior to
CC recovery of starch-derived products. The present sequence encodes
CC alpha-amylase/glucoamylase fusion protein, which is given in the
XX exemplification of the present invention.
XX
SQ Sequence 3285 BP; 573 A; 1399 C; 808 G; 505 T; 0 other;
Query Match 1.8%; Score 44.6; DB 25; Length 3285;
Best Local Similarity 45.2%; Pred. No. 0.66;
Matches 164; Conservative 0; Mismatches 199; Indels 0; Gaps 0;
1838 CGACTCCCGGACTACACTGGGATCGGTGACTCATTCGACACGAGAACATGTCCACGCTGT 1897
780 CGCCTTCAAGTCCACTCCGGTTCGATGGACGACCTCTACACATGATCAACACGCTGAA 839
1898 GGCCTCACTTCGCTACTCTCCACTCTCGAGTATCGTCACTAAGACCTTACAGGTTG 1957
840 GTCCGACTCCCGGACTCCACCTCTCGGACCTTCGTGGAGAACACGACAAACCCGCG 899
1958 GGNAGGCTGACGAGCTCAACAGCCTTTTCGGCAATTCGGCAGCGGGCTCTCTCAA 2017
900 CTTGCGCTCTTACCAACAGATATCGCTTCGCCAAGAACGTGGCGCTTCATCATCT 959
2018 GAATGAGGAGATCTCTGCTTCGCCAGACACTGGCCACCTCTCTCAGAGGTGTCTACCC 2077
960 CACGACGCGATCCCGATCATCTACCGCGCAGGAGACGACATACCGCGGGGACGA 1019
2078 CGCCACTGACAACTTCGCGGCGCGCGTTTCTGCTTCGCGCGAACAATCTGCTCCGT 2137

Db 1020 CCCGCCAACCGCGAGGCCACTCGGTCTCCGGCTACCCGACCGACTCCGAGCTGACAA 1079
QY 2138 GCTGAAGTCGGAGGCAACGTCCTCCATCATCAAGTCCGTTGGCGGAGACTGCGCTCGGCG 2197
Db 1080 GCTCATCGCTCCGCCAACGCCATCCGCACTACGCCATCTCCAGGACACCGGCTTCGT 1139
QY 2198 GGC 2200
Db 1140 GAC 1142
RESULT 14
ABX56041
ID ABX56041 standard; DNA; 1362 BP.
XX
AC ABX56041;
XX
DT 13-FEB-2003 (first entry)
XX
DE M. echinospora calicheamicin biosynthesis gene calQ.
XX
KW Calicheamicin biosynthetic gene cluster; aryltetraaccharide;
KW aglycone; calicheamicin resistance; nonchromoprotein enediynes;
KW enediyne resistance; bone marrow cell; gene therapy; gene; ds.
XX
OS Micromonospora echinospora spp. calichensis.
XX
PN WO200279465-A2.
XX
PD 10-OCT-2002.
XX
PF 28-NOV-2001; 2001WO-US44285.
XX
PR 28-NOV-2000; 2000US-0724797.
XX
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
PI Thorson J;
DR WPI; 2003-092897/08.
P-PSDB; ABU11979.
PT Novel nucleic acid molecule from nonchromoprotein enediyne biosynthetic
PT gene cluster from Micromonospora echinospora useful for conferring
PT calicheamicin resistance on a subject
XX
PS Claim 9; Page 105-107; 179pp; English.
XX
CC The present invention relates to the isolation of the Micromonospora
CC echinospora spp. calichensis calicheamicin biosynthetic gene cluster
CC encoding proteins and enzymes used in the biosynthetic production of
CC calicheamicin, including aryltetraaccharide and aglycone. The gene
CC cluster also includes the gene encoding for the protein that confers
CC calicheamicin resistance. The calicheamicin biosynthetic gene cluster
CC is a nonchromoprotein enediyne biosynthetic gene cluster. Expression
CC vectors containing genes from the biosynthetic gene cluster are also
CC disclosed. The expression vectors are useful for producing calicheamicin
CC biosynthetic proteins. The calicheamicin self-resistance gene provides
CC an approach for gene therapy, for example, by introduction of enediyne
CC resistance genes into bone marrow cells, thus increasing resistance and
CC allowing tolerance to chemotherapeutic doses of calicheamicin.
CC ABX56026-ABX56073 represent genes from the M. echinospora calicheamicin
CC biosynthesis gene cluster.
XX
SQ Sequence 1362 BP; 164 A; 498 C; 521 G; 179 T; 0 other;
Query Match 1.8%; Score 44.4; DB 25; Length 1362;
Best Local Similarity 52.1%; Pred. No. 0.55;
Matches 99; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
2090 CTTCCGCGCGCGCTTCCTTCCTTCGCGCGGCAATGCTGCTCCTCGGTGAGTCGA 2149
Db 186 CCGTCCGCGCGCGCTTCGCGGACCTCCGCGCGGACCGGCTGACGCGGACGACGCTCGTA 245

QY 2150 GCGACGCTCTCATCATCAAGTCGTTGGCGAGACTGCGTGGCGGGCTCAGTCCGG 2209
DB 246 CGACCCGGTCGGCGCGCGGACGTGTGATCTGTGACGGTCGGCACCCCGACGACGCCGG 305
QY 2210 CCTCGGAACTACCGGAGTCTGTAATGAGTGTACCCAGGGAAGATTGCGCGCGTGTCCG 2269
DB 306 CCACGAGATGGTCACCGACCACTCGTCCGGCGTGGAGCAGATCGCCCGCGGTGCG 365
QY 2270 CGCGCGCCGA 2279
DB 366 CGCGCGGCAA 375

RESULT 15
AAS54074
ID AAS54074 standard; DNA; 1221 BP.
AC AAS54074;
XX
DT 13-FEB-2002 (first entry)
XX
DE Pseudomonas aeruginosa DNA for cellular proliferation protein #205.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Pseudomonas aeruginosa.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX
DR WPI; 2001-611495/70.
DR P-PSDB; AAU36215.

XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
PS Claim 27; Seq ID No 7711; 511pp; English.

XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 1221 BP; 214 A; 413 C; 410 G; 184 T; 0 other;

Query Match 1.8%; Score 44.2; DB 23; Length 1221;
Best Local Similarity 48.2%; Pred. No. 0.59;
Matches 124; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 681 GAGTCTGCGCGCGCTCGGCGAGTACTCGAAGATCCCTGACGGCTCGTCAAGTTCTCC 740
DB 136 GATTTGCGCGCGGCATCGCGTGACCTCCCTGGGCCACGGGCATCCGGCGTGTCAAG 195
QY 741 GTGACGCGAGAGATAAGAGAGATCTATAACGAGGAGTCCCGCTCGTCACTGACGTGTCC 800
DB 196 GGGCTCACCGAGCGAGCCCGGCGCATCTGGCACGTTTCCAAGCTTCTCACTAAGCGCG 255
QY 801 GTCCCCCTCGACGCGCGCCAGTGGAGCCTCTCGATTTTCTCCTTTCCGATGTTCAAGACC 860
DB 256 GCGCTGCGCTGGCGCGCAAGCTGGTGACGCCACCTTCGCGGAGCGGTGTTCTCGCC 315
QY 861 GCCTACGTGCGCGTAGCGAAGCTCGAGAACGAGAGATGCTCGACGTTGTCAACGAC 920
DB 316 AATTCCGCGCGCGAAGCAACGAGCGCGCTTCAAGCTGGCTCGCCGCTACGCCAAGAT 375
QY 921 CTCATCGAGTGGCTCAA 937
DB 376 GTCTACGCGCGCGCAGAA 392

Search completed: January 12, 2004, 18:37:59
Job time : 713 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 18:18:51 ; Search time 154 Seconds
(without alignments)
7102.256 Million cell updates/sec

Title: US-09-991-262-47
Perfect score: 2478
Sequence: 1 GTTTTCTTTTACCAAG.....CTGCGTCCCTCGGATACCA 2478

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCUS COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2478	100.0	2478	3	US-08-485-355B-47 Sequence 47, Appl
2	2478	100.0	2478	3	US-08-485-355B-49 Sequence 49, Appl
3	2478	100.0	2478	3	US-09-194-613-1 Sequence 1, Appl
4	2467	99.6	2479	3	US-08-485-355B-51 Sequence 51, Appl
5	71.2	2.9	5312	3	US-08-485-355B-39 Sequence 39, Appl
6	71.2	2.9	5312	3	US-08-485-355B-41 Sequence 41, Appl
7	71.2	2.9	5312	3	US-08-485-355B-43 Sequence 43, Appl
8	71.2	2.9	5368	3	US-08-485-355B-45 Sequence 45, Appl
9	60	2.4	6534	3	US-09-194-613-4 Sequence 4, Appl
10	49	2.0	4403765	3	US-09-103-840A-2 Sequence 2, Appl
11	49	2.0	4411529	3	US-09-103-840A-1 Sequence 1, Appl
12	47.8	1.9	2136	4	US-09-252-991A-11089 Sequence 13089, A
13	47.8	1.9	2406	4	US-09-252-991A-12703 Sequence 12703, A
14	47.8	1.9	2616	4	US-09-252-991A-12851 Sequence 12851, A
15	46.6	1.9	7218	1	US-08-232-463-14 Sequence 14, Appl
16	46	1.9	4403765	3	US-09-103-840A-2 Sequence 2, Appl
17	46	1.9	4411529	3	US-09-103-840A-1 Sequence 1, Appl
18	45.4	1.8	735	3	US-09-003-287-7 Sequence 7, Appl
19	44.2	1.8	972	4	US-09-252-991A-13140 Sequence 13140, A
20	44.2	1.8	972	4	US-09-252-991A-13303 Sequence 13303, A
21	44.2	1.8	2538	4	US-09-252-991A-12659 Sequence 12659, A
22	43.8	1.8	815	3	US-08-818-112-139 Sequence 139, Appl
23	43.8	1.8	815	4	US-08-818-111-134 Sequence 134, Appl
24	43.8	1.8	815	4	US-09-056-556-139 Sequence 139, Appl
25	43.8	1.8	815	4	US-09-072-596-134 Sequence 134, Appl
26	43	1.7	354	4	US-09-252-991A-12550 Sequence 12550, A
27	43	1.7	666	4	US-09-252-991A-6863 Sequence 6863, Ap

28	43	1.7	1164	4	US-09-252-991A-6763	Sequence 6763, Appl
29	41.8	1.7	390	3	US-09-197-649-7	Sequence 7, Appl
30	41.6	1.7	1660	4	US-09-697-367-21	Sequence 21, Appl
31	41.6	1.7	3978	4	US-09-266-965-19	Sequence 19, Appl
32	41.6	1.7	4810	4	US-09-865-621A-3	Sequence 3, Appl
33	41.6	1.7	5163	4	US-09-865-621A-7	Sequence 7, Appl
34	41.6	1.7	5509	4	US-09-865-621A-1	Sequence 1, Appl
35	41.6	1.7	12249	4	US-09-266-965-96	Sequence 96, Appl
36	41.6	1.7	18331	4	US-09-266-965-96	Sequence 96, Appl
37	41	1.7	1767	4	US-09-463-712C-9	Sequence 9, Appl
38	41	1.7	3627	1	US-08-104-072B-6	Sequence 6, Appl
39	41	1.7	3627	1	US-08-351-413-7	Sequence 7, Appl
40	41	1.7	3627	2	US-09-025-583-7	Sequence 7, Appl
41	40.6	1.6	882	3	US-08-818-112-138	Sequence 138, App
42	40.6	1.6	882	4	US-08-818-111-133	Sequence 133, App
43	40.6	1.6	882	4	US-09-056-556-138	Sequence 138, App
44	40.6	1.6	882	4	US-09-072-596-133	Sequence 133, App
45	40.6	1.6	1248	3	US-09-105-537-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-485-355B-47
; Sequence 47, Application US/08485355B
; Patent No. 6177075
; GENERAL INFORMATION:
; APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and Their Uses in
; Protecting Plants
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,355B
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,522
; FILING DATE: 12-MAY-1995
; APPLICATION NUMBER: US 08/089,372
; FILING DATE: 08-JUL-1993
; APPLICATION NUMBER: AU PL4081/92
; FILING DATE: 14-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2478 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 283..753

SEQUENCE DESCRIPTION: SEQ ID NO: 47:									
US-08-485-355B-47									
Query Match 100.0%; Score 2478; DB 3; Length 2478;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GTGTTTCTTTTACCAAGTGGTAAATTTAAACAAGAGAAACACGACCGTAA	60						
DB	1	GTGTTTCTTTTACCAAGTGGTAAATTTAAACAAGAGAAACACGACCGTAA	60						
QY	61	CCGCGCCCTTACACACTCGAGTCCGTGACCAACCGGATTTATAGTCCGCCACACACGGC	120						
DB	61	CCGCGCCCTTACACACTCGAGTCCGTGACCAACCGGATTTATAGTCCGCCACACACGGC	120						
QY	121	GCCTTTTCCGACACTCTCGAGAGTCTGTTGGAGTTTCGTCCGTGACCAACCGGTTGGCA	180						
DB	121	GCCTTTTCCGACACTCTCGAGAGTCTGTTGGAGTTTCGTCCGTGACCAACCGGTTGGCA	180						
QY	181	GTGACAGAGCGTTCGCGACCACTAGAACTCTCGAGCGACGACACACACACACACA	240						
DB	181	GTGACAGAGCGTTCGCGACCACTAGAACTCTCGAGCGACGACACACACACACACA	240						
QY	241	CCGCTTAGTGTGACCTACGGCAGCGTTGATAGCGCGGATTTATAGCGGAGCACACATC	300						
DB	241	CCGCTTAGTGTGACCTACGGCAGCGTTGATAGCGCGGATTTATAGCGGAGCACACATC	300						
QY	301	GCCACTCATCATTAACACCGGTTACACCGGTTACACCGTCCCTTAATACCCCTGAACCTGAA	360						
DB	301	GCCACTCATCATTAACACCGGTTACACCGGTTACACCGTCCCTTAATACCCCTGAACCTGAA	360						
QY	361	GCAGGATGGAGATGTGGAGTGGCTGACAGCGACTCACACCGTCCGCGGAAACCGGTA	420						
DB	361	GCAGGATGGAGATGTGGAGTGGCTGACAGCGACTCACACCGTCCGCGGAAACCGGTA	420						
QY	421	ACGTTCCGGTTCAGCGCAACACCGTCAACCGTCAATGGTAGAAGAAACCAACCGCGTCGGA	480						
DB	421	ACGTTCCGGTTCAGCGCAACACCGTCAACCGTCAATGGTAGAAGAAACCAACCGCGTCGGA	480						
QY	481	CCGGAAGGCAAGTTTCTCCCTGTGACAAATTTACCGCTGTGACAAAGACCTCGCGCAAA	540						
DB	481	CCGGAAGGCAAGTTTCTCCCTGTGACAAATTTACCGCTGTGACAAAGACCTCGCGCAAA	540						
QY	541	GCCTTGACGCAACACCGTCACTTTCCCGCTTAACATCTAGCATGCCGGAATTCGGA	600						
DB	541	GCCTTGACGCAACACCGTCACTTTCCCGCTTAACATCTAGCATGCCGGAATTCGGA	600						
QY	601	ATTGGGCCAAGGAAAGATCGACTCGACTCCGATTCATCCGCTGGTACTTTCAAGTACC	660						
DB	601	ATTGGGCCAAGGAAAGATCGACTCGACTCCGATTCATCCGCTGGTACTTTCAAGTACC	660						
QY	661	TTGACCCAGCGGTGCTACAGAGTCTGCGCGCCCGTCCGCGAGTACTCGAAGATCCCTG	720						
DB	661	TTGACCCAGCGGTGCTACAGAGTCTGCGCGCCCGTCCGCGAGTACTCGAAGATCCCTG	720						
QY	721	ACGGCTCGTCAAGTTCTCGGTGACGACGAGATAGAGAGATCTATAACGAGGATGCC	780						
DB	721	ACGGCTCGTCAAGTTCTCGGTGACGACGAGATAGAGAGATCTATAACGAGGATGCC	780						
QY	781	CCGTGCTGACTGACGTGTCCGTCCTCGACGGCGCCAGTGGAGCCTCTCGATTTTCT	840						
DB	781	CCGTGCTGACTGACGTGTCCGTCCTCGACGGCGCCAGTGGAGCCTCTCGATTTTCT	840						
QY	841	CTTTCCGATGTTTACAGACCGCTACGTCGCGCGTACGAAACGTCGAGAAACAGGAGATGT	900						
DB	841	CTTTCCGATGTTTACAGACCGCTACGTCGCGCGTACGAAACGTCGAGAAACAGGAGATGT	900						
QY	901	CGCTCGACGTGTTCACGACCTCATCGAGTGGCTCAACATCTCGCGGATGCGGTTATG	960						
DB	901	CGCTCGACGTGTTCACGACCTCATCGAGTGGCTCAACATCTCGCGGATGCGGTTATG	960						
QY	961	TCGTTGACTCTGAACAGTGGATTAATTTACCAATGACACCGCTACTAGTCCGCTATC	1020						

DB	961	TGTTGACTCTGAACAGTGGATTAACTTACCNAATGACACACGCTACTACGTCGCGATCC	1020
QY	1021	GGTTCTACGTCCAACTACGAGTTCCAGACCCCCACAGAGGCGCTTGTTCGACAGTCT	1080
DB	1021	GGTTCTACGTCCAACTACGAGTTCCAGACCCCCACAGAGGCGCTTGTTCGACAGTCT	1080
QY	1081	CAGACTACCGGCTCACTTTAAGCGATAACATGTGAAGCCAAACATGCAACACTCGTGG	1140
DB	1081	CAGACTACCGGCTCACTTTAAGCGATAACATGTGAAGCCAAACATGCAACACTCGTGG	1140
QY	1141	ACCAAGGCTTTTGGATCGGCGCCAGTAGCTCTCACCGGACTAGCTACCGGAGTAGG	1200
DB	1141	ACCAAGGCTTTTGGATCGGCGCCAGTAGCTCTCACCGGACTAGCTACCGGAGTAGG	1200
QY	1201	ACGTGAGGAGGCTTACGCTCTGACACTTTTGACCTTCGCGAGACCATTCAGCGCGGTG	1260
DB	1201	ACGTGAGGAGGCTTACGCTCTGACACTTTTGACCTTCGCGAGACCATTCAGCGCGGTG	1260
QY	1261	CACTCGCGTTTGTGGGCGAGTTTGGCCACAGGTTGGCACTGCGCTTCGAGCACTCCAG	1320
DB	1261	CACTCGCGTTTGTGGGCGAGTTTGGCCACAGGTTGGCACTGCGCTTCGAGCACTCCAG	1320
QY	1321	CGTGGAGAGGCACTCTCGGTTGCTACTCACTGCGGCGCAACACGCTACTCTTCC	1380
DB	1321	CGTGGAGAGGCACTCTCGGTTGCTACTCACTGCGGCGCAACACGCTACTCTTCC	1380
QY	1381	CAGCTGGGCTTCGTTAGCTACGTTCTCCTGAGGTTTCCGCTTACGAGCGCTACGACCGA	1440
DB	1381	CAGCTGGGCTTCGTTAGCTACGTTCTCCTGAGGTTTCCGCTTACGAGCGCTACGACCGA	1440
QY	1441	ACGACGGGCTTTGGACCGACTTTGCTTCGCGAGGAGACACCGTCACTTTTCGCGAGGTG	1500
DB	1441	ACGACGGGCTTTGGACCGACTTTGCTTCGCGAGGAGACACCGTCACTTTTCGCGAGGTG	1500
QY	1501	CCGTGCAAGAGTCTGTTGACCAACACCGCGCGCGGCGAGCGCCCGACCTTCA	1560
DB	1501	CCGTGCAAGAGTCTGTTGACCAACACCGCGCGCGGCGAGCGCCCGACCTTCA	1560
QY	1561	CCGTGAGAGTCCCGCTTCAACCGCTTACCAACCGCTTATAGGAAACAGCTCTTAG	1620
DB	1561	CCGTGAGAGTCCCGCTTCAACCGCTTACCAACCGCTTATAGGAAACAGCTCTTAG	1620
QY	1621	AGACTCGACCTCTCTCTGAGGCTCGAACTCCCTATGCCACCTGCTGACTTTGGACAGA	1680
DB	1621	AGACTCGACCTCTCTCTGAGGCTCGAACTCCCTATGCCACCTGCTGACTTTGGACAGA	1680
QY	1681	CGGTGCGCAACAAACCGGAGATCGAGCAGTCTTTTAAAGAAACACTTGGCTGCTATT	1740
DB	1681	CGGTGCGCAACAAACCGGAGATCGAGCAGTCTTTTAAAGAAACACTTGGCTGCTATT	1740
QY	1741	TGGTCCACTCCAAATGCGAAACCCGTTTCCAGCTCAGCCAGCGAGCTCTTTTGGCG	1800
DB	1741	TGGTCCACTCCAAATGCGAAACCCGTTTCCAGCTCAGCCAGCGAGCTCTTTTGGCG	1800
QY	1801	CGGTTTCTTCAACAAATCGGTTATGAGCGCACACCGGACCTCCCGGACTACACTGGCA	1860
DB	1801	CGGTTTCTTCAACAAATCGGTTATGAGCGCACACCGGACCTCCCGGACTACACTGGCA	1860
QY	1861	TCCGTGACTCATTCGACCAAGAAATGTCACCGCTGTGGCGCCACTTCCGCTACTCTCC	1920
DB	1861	TCCGTGACTCATTCGACCAAGAAATGTCACCGCTGTGGCGCCACTTCCGCTACTCTCC	1920
QY	1921	ACTCTGAGTATCGTCACTAGACCTACAGGTTGGGAGGCGCTCAGAACTGCAACA	1980
DB	1921	ACTCTGAGTATCGTCACTAGACCTACAGGTTGGGAGGCGCTCAGAACTGCAACA	1980
QY	1981	CGCTTTTCCGCAATTCGCGCACCGGCGCTCTCAAGAAATGAGGAGATCTCTGCTCG	2040
DB	1981	CGCTTTTCCGCAATTCGCGCACCGGCGCTCTCTCAAGAAATGAGGAGATCTCTGCTCG	2040
QY	2041	CCGACGACTGGCGCACCGGCTCTCAAGGTGTCTACCGCGCTACTGACAACTTCGCGCGG	2100
DB	2041	CCGACGACTGGCGCACCGGCTCTCAAGGTGTCTACCGCGCTACTGACAACTTCGCGCGG	2100

QY 2101 CGTTTCTGCTTCCGCGGAAATGCTCTCTCTCGTGTGAAGTCGGAGCAACGCTCT 2160
 Db 2101 CGTTTCTGCTTCCGCGGAAATGCTCTCTCTCGTGTGAAGTCGGAGCAACGCTCT 2160
 QY 2161 CCATCATCAAGTCGCTTGGGAGACTGCGTGGGCGGCTCAGTCCGCTCGGAGC 2220
 Db 2161 CCATCATCAAGTCGCTTGGGAGACTGCGTGGGCGGCTCAGTCCGCTCGGAGC 2220
 QY 2221 TACCGGAGCTGCTAATGAGTACACAGGAAATGCGCGGCTGTCCGCGCGCGAG 2280
 Db 2221 TACCGGAGCTGCTAATGAGTACACAGGAAATGCGCGGCTGTCCGCGCGCGAG 2280
 QY 2281 CGCGCGCGCGCGCTCTGTCGAATAGTTTGTCTGCTCTCTGTTTGGCGCTTCTGATAA 2340
 Db 2281 CGCGCGCGCGCGCTCTGTCGAATAGTTTGTCTGCTCTCTGTTTGGCGCTTCTGATAA 2340
 QY 2341 AGCGCGTGTCTCGGACATTAACGCTACCTAAGACTCTGCTGAGTCTCCGCTGTACA 2400
 Db 2341 AGCGCGTGTCTCGGACATTAACGCTACCTAAGACTCTGCTGAGTCTCCGCTGTACA 2400
 QY 2401 CGACGGGTCTGCGCGGTTTGGATTCCATTCCCAAGCGGCAAGAGCAGTAGTCTCT 2460
 Db 2401 CGACGGGTCTGCGCGGTTTGGATTCCATTCCCAAGCGGCAAGAGCAGTAGTCTCT 2460
 QY 2461 GGTCTCTCGGATACCA 2478
 Db 2461 GGTCTCTCGGATACCA 2478

RESULT 2

US-08-485-355B-49
 ; Sequence 49, Application US/08485355B
 ; Patent No. 6177075

GENERAL INFORMATION:

; APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
 ; TITLE OF INVENTION: Insect Viruses and Their Uses in
 ; Protecting Plants

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
 ; STREET: Four Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States
 ; ZIP: 94111-4187

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/485,355B
 ; FILING DATE: 07-Jun-1995

CLASSIFICATION DATA:

; APPLICATION NUMBER: US 08/440,522
 ; FILING DATE: 12-MAY-1995

; APPLICATION NUMBER: US 08/089,372
 ; FILING DATE: 08-JUL-1993

; APPLICATION NUMBER: AU PL4081/92
 ; FILING DATE: 14-AUG-1992

ATTORNEY/AGENT INFORMATION:

; NAME: Trecartin, Richard P.
 ; REGISTRATION NUMBER: 31,801
 ; REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 398-3249
 ; TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 49:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2478 base pairs

TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 366..2306
 SEQUENCE DESCRIPTION: SEQ ID NO: 49:
 US-08-485-355B-49

Query Match 100.0%; Score 2478; DB 3; Length 2478;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTTTCTTTCTTTACCAAGTGTGGTAAATTTAAACAAAGAAAGAAACCGAGCGTAA 60
 Db 1 GTTTTCTTTCTTTTACCAAGTGTGGTAAATTTAAACAAAGAAAGAAACCGAGCGTAA 60

QY 61 CCCGCCCTTACACACCTCGAGTCCGTGACACCGGATTATACGTGCGCCACCAACGCG 120
 Db 61 CCCGCCCTTACACACCTCGAGTCCGTGACACCGGATTATACGTGCGCCACCAACGCG 120

QY 121 GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTTCTCGTGCACACCGGTTGGCA 180
 Db 121 GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTTCTCGTGCACACCGGTTGGCA 180

QY 181 GTCGACAGACGCTTCCGGACCACTAGAACCTCTCGAGCGACGACACACACACACA 240
 Db 181 GTCGACAGACGCTTCCGGACCACTAGAACCTCTCGAGCGACGACACACACACACA 240

QY 241 CGCCTTAGTGTGACCTACCGCAGCGTTGATAGCGCGGATTATGAGGAGACACCATC 300
 Db 241 CGCCTTAGTGTGACCTACCGCAGCGTTGATAGCGCGGATTATGAGGAGACACCATC 300

QY 301 GCCCACTCCATCAGATTACACCGGTTTACACCTTCCCTTAATACCCCTGAACCTGAA 360
 Db 301 GCCCACTCCATCAGATTACACCGGTTTACACCTTCCCTTAATACCCCTGAACCTGAA 360

QY 361 GCAGGATGGAGATGCTGGAGTGGCTCACAGCGACCTCACAAACCGTCGCGAAACCGTA 420
 Db 361 GCAGGATGGAGATGCTGGAGTGGCTCACAGCGACCTCACAAACCGTCGCGAAACCGTA 420

QY 421 ACCTTTCGGGTGAGCGCCAAACACCGTCAACGTTAGTAGAAGAAACCAACCGGCTCGGA 480
 Db 421 ACCTTTCGGGTGAGCGCCAAACACCGTCAACGTTAGTAGAAGAAACCAACCGGCTCGGA 480

QY 481 CCGGAAGGCAAGTTTCTCCCTTGCACCAATTTACCGCTGCTGCAACAGACCTCGCGCAA 540
 Db 481 CCGGAAGGCAAGTTTCTCCCTTGCACCAATTTACCGCTGCTGCAACAGACCTCGCGCAA 540

QY 541 GCCTTGACGCCAAACACCGTCACTTTCCCGGCTAAACATCTAGCATGCCGAAATTCGGA 600
 Db 541 GCCTTGACGCCAAACACCGTCACTTTCCCGGCTAAACATCTAGCATGCCGAAATTCGGA 600

QY 601 ATTGGCCAAAGGAAAGATCGACCTCGATTCGATTCCATCGGCTGGTACTTCAAGTACC 660
 Db 601 ATTGGCCAAAGGAAAGATCGACCTCGATTCGATTCCATCGGCTGGTACTTCAAGTACC 660

QY 661 TTGACCCAGCGGTGCTACAGAGTCTGCGCGCGCGTGGCGAGTACTCGAGATCCCTG 720
 Db 661 TTGACCCAGCGGTGCTACAGAGTCTGCGCGCGCGTGGCGAGTACTCGAGATCCCTG 720

QY 721 ACGGCTCGTCAAGTTTCTCCGTCGACGAGATGAAGAGATCTATAACGAGGAGTGC 780
 Db 721 ACGGCTCGTCAAGTTTCTCCGTCGACGAGATGAAGAGATCTATAACGAGGAGTGC 780

QY 781 CCGTCTCACTGACGTTGCTCCCTCGAGCGCGCGAGTGGAGCTCTCGATTTCT 840
 Db 781 CCGTCTCACTGACGTTGCTCCCTCGAGCGCGCGAGTGGAGCTCTCGATTTCT 840

QY 841 CTTTTCGATGTTAGAACCGGCTACGTGCGCGTACGAAACGTCGAGAACGAGGATGT 900
 Db 841 CTTTTCGATGTTTTCAGAACCGGCTACGTGCGCGTACGAAACGTCGAGAACGAGGATGT 900

; STRANDEDNESS: double		Query Match		100.0%; Score 2478; DB 3; Length 2478;	
; TOPOLOGY: linear		Best Local Similarity		100.0%; Pred. No. 0;	
; MOLECULE TYPE: cDNA		Matches 2478; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
US-09-194-613-1					
Qy	1	GTGTTTCTTTCTTTACCAAGTGTGTAATAATTTAAACAAAGAGAAACACGAGACCGTAA	60		
Db	1	GTGTTTCTTTCTTTACCAAGTGTGTAATAATTTAAACAAAGAGAAACACGAGACCGTAA	60		
Qy	61	CCGGCCCTTACACACTCGAGTCCGTGACACCCGGAATTATACGTGCGCCACACACGCG	120		
Db	61	CCGGCCCTTACACACTCGAGTCCGTGACACCCGGAATTATACGTGCGCCACACACGCG	120		
Qy	121	GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTTCGTGCGTGACACCCGGTTGGCA	180		
Db	121	GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTTCGTGCGTGACACCCGGTTGGCA	180		
Qy	181	GTGACAGACGCTTCCGGACCACTAGAACTCTCTCGAGCGACGACACACACACACA	240		
Db	181	GTGACAGACGCTTCCGGACCACTAGAACTCTCTCGAGCGACGACACACACACACA	240		
Qy	241	CCGCTTAGCTGCACTTACCGGAGCGTTGATAGCGCGGATTTATGAGGAGCACACCATC	300		
Db	241	CCGCTTAGCTGCACTTACCGGAGCGTTGATAGCGCGGATTTATGAGGAGCACACCATC	300		
Qy	301	GCCCACTCCATACATTACCAACCCGTTACACCTTGCCCTAATACCCCTCAACCTGAA	360		
Db	301	GCCCACTCCATACATTACCAACCCGTTACACCTTGCCCTAATACCCCTCAACCTGAA	360		
Qy	361	GCAGATGGGAGATGCTGGAGTGGCTCACAGCGACTCACAAACCGTCCGCGAAACCCGTA	420		
Db	361	GCAGATGGGAGATGCTGGAGTGGCTCACAGCGACTCACAAACCGTCCGCGAAACCCGTA	420		
Qy	421	ACGTTCCGGTTCAGCGCCAAACCGTCAACGTTCAATGGTAGAAGAAACCAACCGGCTCGGA	480		
Db	421	ACGTTCCGGTTCAGCGCCAAACCGTCAACGTTCAATGGTAGAAGAAACCAACCGGCTCGGA	480		
Qy	481	CCGGAAGCAAGTTTCTCCCTGCAATTTACCGCTGCTGCAAGACCTTCGCGCAAA	540		
Db	481	CCGGAAGCAAGTTTCTCCCTGCAATTTACCGCTGCTGCAAGACCTTCGCGCAAA	540		
Qy	541	GCCTTGACGCCAAACACCGTCACTTTCCCGCTAAACATCTCTAGCATGCGCGAATTCGCGA	600		
Db	541	GCCTTGACGCCAAACACCGTCACTTTCCCGCTAAACATCTCTAGCATGCGCGAATTCGCGA	600		
Qy	601	ATTGGGCCAAGGAAAGATCGACCTCGACTCCGATTCCATCGGCTGGTACTTCAAGTACC	660		
Db	601	ATTGGGCCAAGGAAAGATCGACCTCGACTCCGATTCCATCGGCTGGTACTTCAAGTACC	660		
Qy	661	TTGACCCAGCGGTGCTACAGAGTCTGCGCGCGCGTCCGCGAGTACTCGAAGATCCCTTG	720		
Db	661	TTGACCCAGCGGTGCTACAGAGTCTGCGCGCGCGTCCGCGAGTACTCGAAGATCCCTTG	720		
Qy	721	ACGGCTCTCAAGTTCTCCGTGACGAGATAGAGATCTATACAGGAGATGCC	780		
Db	721	ACGGCTCTCAAGTTCTCCGTGACGAGATAGAGATCTATACAGGAGATGCC	780		
Qy	781	CCGTGCTCACTGACGTGCTCCGTCGCCCTCGACGGCCGCGAGTGGAGCTCTCGATTTTCT	840		
Db	781	CCGTGCTCACTGACGTGCTCCGTCGCCCTCGACGGCCGCGAGTGGAGCTCTCGATTTTCT	840		
Qy	841	CTTTCCGATGTTACAGAACCGCTTACGTCCGCGTAGCGAACTCGAGAACAGGAGATGT	900		
Db	841	CTTTCCGATGTTACAGAACCGCTTACGTCCGCGTAGCGAACTCGAGAACAGGAGATGT	900		
Qy	901	CGCTCGACGTTGTCAAGACCTCATCGAGTGGCTCAACATCTCCGCGACTCGCGTTATG	960		
Db	901	CGCTCGACGTTGTCAAGACCTCATCGAGTGGCTCAACATCTCCGCGACTCGCGTTATG	960		

QY	961	TCGTTGACTCTGAAACAGTGGATTAACTTCAACATGACACACGACTACTGTCGCGATCC	1020		
DB	961	TCGTTGACTCTGAAACAGTGGATTAACTTCAACATGACACACGACTACTGTCGCGATCC	1020		
QY	1021	GGTTTCTACGTCCCAACCTACGAGTTTCCAGACCCACACAGAGGGCTTGTTCGACAGTCT	1080		
DB	1021	GGTTTCTACGTCCCAACCTACGAGTTTCCAGACCCACACAGAGGGCTTGTTCGACAGTCT	1080		
QY	1081	CAGACTACCGCTCACTTATAGGCGATAACATGTGAAGCCAAACATGCCAACCTCGTGG	1140		
DB	1081	CAGACTACCGCTCACTTATAGGCGATAACATGTGAAGCCAAACATGCCAACCTCGTGG	1140		
QY	1141	ACCAAGGCTTTTGGATCGGCGCCAGTACGTCTCACCCCGACTAGCTACCGAGTAGC	1200		
DB	1141	ACCAAGGCTTTTGGATCGGCGCCAGTACGTCTCACCCCGACTAGCTACCGAGTAGC	1200		
QY	1201	ACGTTCAGCGAGGCTTACGTCTGACACTTTTGACCTTCCGAGAGCATCCAGCGCGCTG	1260		
DB	1201	ACGTTCAGCGAGGCTTACGTCTGACACTTTTGACCTTCCGAGAGCATCCAGCGCGCTG	1260		
QY	1261	CACTCGCGTTTGTGGGCGAGGTTTGGCCACAGGGTGGCACTGCGGCTTGAGGCACTCCAG	1320		
DB	1261	CACTCGCGTTTGTGGGCGAGGTTTGGCCACAGGGTGGCACTGCGGCTTGAGGCACTCCAG	1320		
QY	1321	CCTGGAGAGGCGATCTCTCGGTTGCTACTCACTCGGCGCCACAAACGGTACTACTTTCC	1380		
DB	1321	CCTGGAGAGGCGATCTCTCGGTTGCTACTCACTCGGCGCCACAAACGGTACTACTTTCC	1380		
QY	1381	CAGCTGGCTTCGGTTAGCTACGTTCTCCCTGAGGGTTTCGCGCTTGAGGCGTACGACCGA	1440		
DB	1381	CAGCTGGCTTCGGTTAGCTACGTTCTCCCTGAGGGTTTCGCGCTTGAGGCGTACGACCGA	1440		
QY	1441	ACGACGGCTCTTTGGACCGACTTTCGCTCCGAGAGAGACACCGTCACTTTCCGCGAGGTCG	1500		
DB	1441	ACGACGGCTCTTTGGACCGACTTTCGCTCCGAGAGAGACACCGTCACTTTCCGCGAGGTCG	1500		
QY	1501	CCGTGCGAGAGTCTGTTGACCAACACCCCGCGCGCGGCGAGCGCCCCACCTTCA	1560		
DB	1501	CCGTGCGAGAGTCTGTTGACCAACACCCCGCGCGCGGCGAGCGCCCCACCTTCA	1560		
QY	1561	CCGTGAGAGTCCCGCTTCAACCGCTTACCAACACCGCTTGTAGGACAGCTCTTAG	1620		
DB	1561	CCGTGAGAGTCCCGCTTCAACCGCTTACCAACACCGCTTGTAGGACAGCTCTTAG	1620		
QY	1621	AGACTCGACCTCTCTCTGTTAGCTGAACTCCCTATGCTCACTGCTGACTTTGGACAGA	1680		
DB	1621	AGACTCGACCTCTCTCTGTTAGCTGAACTCCCTATGCTCACTGCTGACTTTGGACAGA	1680		
QY	1681	CGGTTCGCAACAAACCCGAAATCGAGCAGTCCGCTTTTAAAGAAACACTTGGCTGCTATT	1740		
DB	1681	CGGTTCGCAACAAACCCGAAATCGAGCAGTCCGCTTTTAAAGAAACACTTGGCTGCTATT	1740		
QY	1741	TGGTTCACCTCAAAATGCGAACCCGCTTTCAGCTCACGCGAGCGAGCTCTTTGGG	1800		
DB	1741	TGGTTCACCTCAAAATGCGAACCCGCTTTCAGCTCACGCGAGCGAGCTCTTTGGG	1800		
QY	1801	CGGTTCCTTCAACAAATCGGCTTATGAGCGCACACGCGACCTCCCGGACTACCTGGCA	1860		
DB	1801	CGGTTCCTTCAACAAATCGGCTTATGAGCGCACACGCGACCTCCCGGACTACCTGGCA	1860		
QY	1861	TCCGTGACTCATTCGACCAAGATGTCCACCGCTGTGGCCACTTCCGCTCACTCTCC	1920		
DB	1861	TCCGTGACTCATTCGACCAAGATGTCCACCGCTGTGGCCACTTCCGCTCACTCTCC	1920		
QY	1921	ACTCTGCACTCATTCGACCAAGATGTCCACCGCTGTGGCCACTTCCGCTCACTCTCC	1980		
DB	1921	ACTCTGCACTCATTCGACCAAGATGTCCACCGCTGTGGCCACTTCCGCTCACTCTCC	1980		
QY	1981	CGCTTTTCCGCGCAATTCGCGACCGGGCTCTCTCAAGATGAGGAGATCTCTGCTCG	2040		
DB	1981	CGCTTTTCCGCGCAATTCGCGACCGGGCTCTCTCAAGATGAGGAGATCTCTGCTCG	2040		
QY	2041	CCGACGACTGGCCACCCGCTCTCACAGGTGTCTACCCCGCCACTGACAACTTCCGCGCG	2100		

Db 2041 CCGAGACCTGGCCACCGCTCTCACAGTGTCTACCCCGCCACTGACAACTTCGGGCGG 2100
QY 2101 CGGTTTCCTCTCGCGGGAACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
Db 2101 CGGTTTCCTCTCGCGGGAACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
QY 2161 CCATCATCAAGTCGCTGCGGAGACTGCGTCTGCGGCGGCTCAGTCCGCGCTCGGAAGC 2220
Db 2161 CCATCATCAAGTCGCTGCGGAGACTGCGTCTGCGGCGGCTCAGTCCGCGCTCGGAAGC 2220
QY 2221 TACCCGGAAGTCTGTAAGTGTACCAAGGAAGATTGCGCGGCTGTCGCGGCGCGCGAG 2280
Db 2221 TACCCGGAAGTCTGTAAGTGTACCAAGGAAGATTGCGCGGCTGTCGCGGCGCGCGAG 2280
QY 2281 CGGCGCGCGCGCGCTGTCGCAATTAGTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 2340
Db 2281 CGGCGCGCGCGCGCTGTCGCAATTAGTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 2340
QY 2341 AGGCGGTGTCGCGGACATTCAGCGTACCTAAAGACTCTGTGAGTCCCGCTCGTTACA 2400
Db 2341 AGGCGGTGTCGCGGACATTCAGCGTACCTAAAGACTCTGTGAGTCCCGCTCGTTACA 2400
QY 2401 CGACGGGTGTCGCGGCTGCTGATTCATTCCTCAAGCGGCAAGAGACGTAGTTAGCTCT 2460
Db 2401 CGACGGGTGTCGCGGCTGCTGATTCATTCCTCAAGCGGCAAGAGACGTAGTTAGCTCT 2460
QY 2461 GGTCTCTCGGATACCA 2478
Db 2461 GGTCTCTCGGATACCA 2478

RESULT 4

US-08-485-355B-51
; Sequence 51, Application US/0848535B
; Patent No. 6177075
; GENERAL INFORMATION:
; APPLICANT: Christian, P. D., Gordon, K. H. J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and Their Uses in
; Protecting Plants
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,355B
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,522
; FILING DATE: 12-MAY-1995
; APPLICATION NUMBER: US 08/089,372
; FILING DATE: 08-JUL-1993
; APPLICATION NUMBER: AU PL4081/92
; FILING DATE: 14-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:
; LENGTH: 2479 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 283..2307
; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-08-485-355B-51

Query Match 99.6%; Score 2467; DB 3; Length 2479;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2478; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GTTTTCTTTCTTTACCAAGTGTGTTAAATTTAAACAAGAAAGAAACCCAGGACCGTAA 60
Db 1 GTTTTCTTTCTTTACCAAGTGTGTTAAATTTAAACAAGAAAGAAACCCAGGACCGTAA 60

QY 61 CCGCGCCCTTACACACCTCGAGTCCGTGACACCGGATTATACGTGCGCCACACACCGC 120
Db 61 CCGCGCCCTTACACACCTCGAGTCCGTGACACCGGATTATACGTGCGCCACACACCGC 120

QY 121 GCCTTTTCCGACCACTCTCGAGAGTCGTTGGAGTTTCGTCGTGACCAACCGGTTGGCA 180
Db 121 GCCTTTTCCGACCACTCTCGAGAGTCGTTGGAGTTTCGTCGTGACCAACCGGTTGGCA 180

QY 181 GTGACAGACGCTTCGGACCACTAGAACCTCTCTGAGGAGCGGACACACAGCACACA 240
Db 181 GTGACAGACGCTTCGGACCACTAGAACCTCTCTGAGGAGCGGACACACAGCACACA 240

QY 241 CGCCTTAGCTGACCTACGCGAGGTTGATAGCGCGGATTATGAGCGAGACACACCATC 300
Db 241 CGCCTTAGCTGACCTACGCGAGGTTGATAGCGCGGATTATGAGCGAGACACACCATC 300

QY 301 GCCCACTCCATCACATTACCAACCGGTTAGACCCCTTGCCCTTAATACCCCTGAAC 360
Db 301 GCCCACTCCATCACATTACCAACCGGTTAGACCCCTTGCCCTTAATACCCCTGAAC 360

QY 361 GCAGGATGGAGATCTGAGTGGGCTCACAGGACCTCACACCGTTCGCGGAAACCGTGA 420
Db 361 GCAGGATGGAGATCTGAGTGGGCTCACAGGACCTCACACCGTTCGCGGAAACCGTGA 420

QY 421 AGCTTCGGGTGACGCGCAACACCGTCACCTCAATGTTAGAGAAACCAACCGCGTCGA 480
Db 421 AGCTTCGGGTGACGCGCAACACCGTCACCTCAATGTTAGAGAAACCAACCGCGTCGA 480

QY 481 CCGGAAGGCAAGTTTCTCCCTCGACAAATTTCCCGCTGCTGCAACAAGACCTTCGCGCAA 540
Db 481 CCGGAAGGCAAGTTTCTCCCTCGACAAATTTCCCGCTGCTGCAACAAGACCTTCGCGCAA 540

QY 541 GCCTTGACGCCAACACCGTCACTTT-CCCGCTTAACATCTCTAGCATGCCGAATTCGG 599
Db 541 GCCTTGACGCCAACACCGTCACTTT-CCCGCTTAACATCTCTAGCATGCCGAATTCGG 600

QY 600 AATTGGGCCAAGGGAAGATCGACCTCGACTCCGATTCCATCGGTGGTACTTCAAGTAC 659
Db 601 AATTGGGCCAAGGGAAGATCGACCTCGACTCCGATTCCATCGGTGGTACTTCAAGTAC 660

QY 660 CTTGACCCAGCGGGTGTACAGAGTCTGCGCGCGCGCTGCGGAGTACTCTGAAGATCCCT 719
Db 661 CTTGACCCAGCGGGTGTACAGAGTCTGCGCGCGCGCTGCGGAGTACTCTGAAGATCCCT 720

QY 720 GACGCGCTCGTCAAGTTCTCCCTCGACCGGAGATAAGAGATCTATACGAGGAGTGC 779
Db 721 GACGCGCTCGTCAAGTTCTCCCTCGACCGGAGATAAGAGATCTATACGAGGAGTGC 780

QY 780 CCGGTCGTCACGTGAGTGTCCCTCCCTCGACGCGCGCGCTGCGGAGTACTCTCGATTTTC 839
Db 781 CCGGTCGTCACGTGAGTGTCCCTCCCTCGACGCGCGCGCTGCGGAGTACTCTCGATTTTC 840

QY 840 TCCTTTCCGATGTTTCAAGAACCGCCTACGTCGCGGTAGCGAAACGTCGAGAGATG 899

Db 841 TCCCTTCGATGTTTCAAGAACCGCTACGTGCGCGTAGCGAAGCGTTCGAGAACAGGAGATG 900
QY TCGCTCGAGTGTGTCAGACACCTCATCGAGTGCGTCAACAATCTCGCCGACTGGCGTTAT 959
Db TCGCTCGAGTGTGTCAGACACCTCATCGAGTGCGTCAACAATCTCGCCGACTGGCGTTAT 960
QY GTGCTTGACTGTGAACAGTGGATTAATCTTCAACAATGAACAACAGTACTCGTTCGCATC 1019
Db GTGCTTGACTGTGAACAGTGGATTAATCTTCAACAATGAACAACAGTACTCGTTCGCATC 1020
QY CGGCTTCTAGCTTCAACCTACGAGTTCAGAGTTCAGAGCCCAAGAGGGCGCTTGTTCGACAGTC 1079
Db CGGCTTCTAGCTTCAACCTACGAGTTCAGAGTTCAGAGCCCAAGAGGGCGCTTGTTCGACAGTC 1080
QY TCAGACTACCGCTCACTTATAGGCGATAAATGTGAAGCCAAATGCAACACTTCGTC 1139
Db TCAGACTACCGCTCACTTATAGGCGATAAATGTGAAGCCAAATGCAACACTTCGTC 1140
QY GACCAAGGCTTTTGGATTCGCGCGCAGTAGCTCTCACCCGACTAGCTTACCGCAGTAC 1199
Db GACCAAGGCTTTTGGATTCGCGCGCAGTAGCTCTCACCCGACTAGCTTACCGCAGTAC 1200
QY GAGCTCAGCGAGGCTTACGCTCTGACACTTTGACCTTCGCCAGACCATCCAGCGCGCT 1259
Db GAGCTCAGCGAGGCTTACGCTCTGACACTTTGACCTTCGCCAGACCATCCAGCGCGCT 1260
QY GCACTCGCTTTGTGCGGAGGTTTCCACAGGCTGGCACTTCGCGCTCGAGGCACTCCA 1319
Db GCACTCGCTTTGTGCGGAGGTTTCCACAGGCTGGCACTTCGCGCTCGAGGCACTCCA 1320
QY GCTGGGAGCAGGATCTCTCGGTGCTACTCACCTCGCGCCACAAACGCTACTCTTTC 1379
Db GCTGGGAGCAGGATCTCTCGGTGCTACTCACCTCGCGCCACAAACGCTACTCTTTC 1380
QY CAGCTCGCTTCGTTAGCTACGTTCTCCTGAGGTTTTCGCTTTCGAGGCTTTCGAGCCG 1439
Db CAGCTCGCTTCGTTAGCTACGTTCTCCTGAGGTTTTCGCTTTCGAGGCTTTCGAGCCG 1440
QY AACGACGGCTTTTGGACCGACTTCGCTTCGCGAGGAGACCGTCACTTTTCGCGAGGTC 1499
Db AACGACGGCTTTTGGACCGACTTCGCTTCGCGAGGAGACCGTCACTTTTCGCGAGGTC 1500
QY GCGCTCGAGAGTCTGTGACCAACACCCCGCGCGCGGCGAGCGCCCGACCTTC 1559
Db GCGCTCGAGAGTCTGTGACCAACACCCCGCGCGCGGCGAGCGCCCGACCTTC 1560
QY ACCGTGAGAGTGCCCGCTTCAACACCGCTTACCAACACCGGTGTTTGAAGAACACGCTCTTA 1619
Db ACCGTGAGAGTGCCCGCTTCAACACCGCTTACCAACACCGGTGTTTGAAGAACACGCTCTTA 1620
QY GAGACTCGACCTCTCTGCTAGGCTCGACTCCCTTATGCCACCTGCTGACTTTGACAG 1679
Db GAGACTCGACCTCTCTGCTAGGCTCGACTCCCTTATGCCACCTGCTGACTTTGACAG 1680
QY ACCGTGCGCAACACCCGAGTTCGAGAGTCTGCTTCTTAAAGAAACACTTGGCTCTAT 1739
Db ACCGTGCGCAACACCCGAGTTCGAGAGTCTGCTTCTTAAAGAAACACTTGGCTCTAT 1740
QY TTGGTCCACTTCCAAATTCGAAACCCCGTTTCCAGCTCAAGCCAGCTCTTTCG 1799
Db TTGGTCCACTTCCAAATTCGAAACCCCGTTTCCAGCTCAAGCCAGCTCTTTCG 1800
QY GCGCTTCTTCAACACTCGGCTTATGAGCGCACGCGACTCCCGGACTACTGGC 1859
Db GCGCTTCTTCAACACTCGGCTTATGAGCGCACGCGACTCCCGGACTACTGGC 1860
QY ATCCGTGACTCTATTCGACAGAACATGTCCACCGCTGTGCGCCACTTTCGCTCACTCTCC 1919
Db ATCCGTGACTCTATTCGACAGAACATGTCCACCGCTGTGCGCCACTTTCGCTCACTCTCC 1920
QY CACTCTCGAGTATCGTCACTAAGACCTACCGGGTTGGAGGGCTTCAGAACGTCACAC 1979

Db 1921 CACTCTCGAGTATCGTCACTAAGACCTACACAGGGTTGGAGGGCTCAGAACGTCACAC 1980
QY AGCCCTTTTCGCGCAATTCGCGCACGCGGGCTCTCTCAAGAAATGAGAGATCTCTGCGCTC 2039
Db AGCCCTTTTCGCGCAATTCGCGCACGCGGGCTCTCTCAAGAAATGAGAGATCTCTGCGCTC 2040
QY GCGGAGGACTGCGCCACCGCTCTCAGAGTGTCTACCGGCGCACATGACAACTTCGCGGCC 2099
Db GCGGAGGACTGCGCCACCGCTCTCAGAGTGTCTACCGGCGCACATGACAACTTCGCGGCC 2100
QY GCGGTTTCTGCTTTCGCGCGCAACATGCTCTCGCTGCTGAAGTTCGAGGCAACGCTC 2159
Db GCGGTTTCTGCTTTCGCGCGCAACATGCTCTCGCTGCTGAAGTTCGAGGCAACGCTC 2160
QY TCATCATCAAGTTCGCGGAGACTGCGCTCGCGCGGCTCAGTTCGCGGCTTCGCGAAG 2219
Db TCATCATCAAGTTCGCGGAGACTGCGCTCGCGCGGCTCAGTTCGCGGCTTCGCGAAG 2220
QY CTACCGGAGTCTAATGAGTGTACAGGGAAGATTCGCGCGGCTTCGCGCGGCGCGCA 2279
Db CTACCGGAGTCTAATGAGTGTACAGGGAAGATTCGCGCGGCTTCGCGCGGCGCGCA 2280
QY GCGCGCGCGCGCGCTCGTCCCAATAGTTTCTCGCTCTCTGTTTCGCGGTTTCGTA 2339
Db GCGCGCGCGCGCGCTCGTCCCAATAGTTTCTCGCTCTCTGTTTCGCGGTTTCGTA 2340
QY AACGCGGCTGCTCCCGCACATTCAGCGGTACCTTAAAGACTCTGCTGAGTCCCGCTTCGTTAC 2399
Db AACGCGGCTGCTCCCGCACATTCAGCGGTACCTTAAAGACTCTGCTGAGTCCCGCTTCGTTAC 2400
QY ACGAGGCTTCGCGCGGTTTCGATTCATTCCTCAAGCGGCAAGAGGAGTGTAGTCTC 2459
Db ACGAGGCTTCGCGCGGTTTCGATTCATTCCTCAAGCGGCAAGAGGAGTGTAGTCTC 2460
QY TCGCTTCCTTCGCGGATACCA 2478
Db TCGCTTCCTTCGCGGATACCA 2479

RESULT 5

US-08-485-355B-39
; Sequence 39, Application US/08485355B
; Patent No. 6177075
; GENERAL INFORMATION:
; APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and Their Uses in .
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,355B
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,522
; FILING DATE: 12-MAY-1995
; APPLICATION NUMBER: US 08/089,372
; FILING DATE: 08-JUL-1993
; APPLICATION NUMBER: AU PL4081/92
; FILING DATE: 14-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.

REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:
LENGTH: 5312 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:

NAME/KEY: CDS

LOCATION: 37..5148

SEQUENCE DESCRIPTION: SEQ ID NO: 39:

US-08-485-355B-39

Query Match 2.9%; Score 71.2; DB 3; Length 5312;
Best Local Similarity 78.7%; Pred. No. 5.5e-08;
Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 2371 TAAAGACTCTGGTGGAGTCCCGCTGTACACGACGGGTCTGCGCGGTTTCGATTCCATTC 2430
|||||
DB 5205 TAAAGACTCTGGTGGAGTCTCAACGTTACTCGTTGAGTCTGCTGCGGTTTCGATTCCATTC 5264

QY 2431 CCAAGCGGCAAGAAGACGTAGTCTGCTGCGTCCCTCGGATACCA 2478
|||||
DB 5265 CCAAGCAGCAAAAGGTGCGCAACTAGTACGGCGGCCCTCGGATACCA 5312

RESULT 6

US-08-485-355B-41
Sequence 41, Application US/08485355B
Patent No. 6177075

GENERAL INFORMATION:

APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
TITLE OF INVENTION: Insect Viruses and Their Uses in
Protecting Plants

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
STREET: Four Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: United States

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,355B

FILING DATE: 07-Jun-1995

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/440,522

FILING DATE: 12-MAY-1995

APPLICATION NUMBER: US 08/089,372

FILING DATE: 08-JUL-1993

APPLICATION NUMBER: AU PL4081/92

FILING DATE: 14-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: Trecartin, Richard F.

REGISTRATION NUMBER: 31,801

REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:
LENGTH: 5312 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 4218..4514
SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-08-485-355B-41

Query Match 2.9%; Score 71.2; DB 3; Length 5312;
Best Local Similarity 78.7%; Pred. No. 5.5e-08;
Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 2371 TAAAGACTCTGGTGGAGTCCCGCTGTACACGACGGGTCTGCGCGGTTTCGATTCCATTC 2430
|||||
DB 5205 TAAAGACTCTGGTGGAGTCTCAACGTTACTCGTTGAGTCTGCTGCGGTTTCGATTCCATTC 5264

QY 2431 CCAAGCGGCAAGAAGACGTAGTCTGCTGCGTCCCTCGGATACCA 2478
|||||
DB 5265 CCAAGCAGCAAAAGGTGCGCAACTAGTACGGCGGCCCTCGGATACCA 5312

RESULT 7

US-08-485-355B-43

Sequence 43, Application US/08485355B

Patent No. 6177075

GENERAL INFORMATION:

APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
TITLE OF INVENTION: Insect Viruses and Their Uses in
Protecting Plants

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP

STREET: Four Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: United States

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,355B

FILING DATE: 07-Jun-1995

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/440,522

FILING DATE: 12-MAY-1995

APPLICATION NUMBER: US 08/089,372

FILING DATE: 08-JUL-1993

APPLICATION NUMBER: AU PL4081/92

FILING DATE: 14-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: Trecartin, Richard F.

REGISTRATION NUMBER: 31,801

REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 5312 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: CDS
LOCATION: 4518..4937
SEQUENCE DESCRIPTION: SEQ ID NO: 43;
US-08-485-355B-43

Query Match 2.9%; Score 71.2; DB 3; Length 5312;
Best Local Similarity 78.7%; Pred. No. 5.5e-08;
Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 2371 TAAAGACTCTGCTGAGTCCCGCTTACACACGGGCTCGCGGGTTTCGATTCCATTC 2430
Db TAAAGACTCTGCTGAGTCTCAACGTTACTCGTTGAGTCTGCTGGGTTTCGATTCCATTC 5264

QY 2431 CCAAGCGGCAAGAGACCTAGTCTGCTGCTCGGATACCA 2478
Db CCAAGCAGCAAGAGGTCGCACTAGTACGGCGCCCTGGGATACCA 5312

RESULT 8
US-08-485-355B-45
; Sequence 45, Application US/08485355B
; Patent No. 6177075
; GENERAL INFORMATION:
; APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and their Uses in
; Protecting Plants
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,355B
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,522
; FILING DATE: 12-MAY-1995
; APPLICATION NUMBER: US 08/089,372
; FILING DATE: 08-JUL-1993
; APPLICATION NUMBER: AU PL4081/92
; FILING DATE: 14-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Treccartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5368 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4944..5162
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-485-355B-45

Query Match 2.9%; Score 71.2; DB 3; Length 5368;
Best Local Similarity 78.7%; Pred. No. 5.5e-08;

Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 2371 TAAAGACTCTGCTGAGTCCCGCTTACACACGGGCTCGCGGGTTTCGATTCCATTC 2430
Db TAAAGACTCTGCTGAGTCTCAACGTTACTCGTTGAGTCTGCTGGGTTTCGATTCCATTC 5320

QY 2431 CCAAGCGGCAAGAGGACCTAGTCTGCTGCTCGGATACCA 2478
Db CCAAGCAGCAAGAGGTCGCACTAGTACGGCGCCCTGGGATACCA 5368

RESULT 9
US-09-194-613-4
; Sequence 4, Application US/09194613
; Patent No. 6251654
; GENERAL INFORMATION:
; APPLICANT: GORDON, Karl H.
; APPLICANT: HANZLIK, Terry N.
; TITLE OF INVENTION: MODIFIED SMALL RNA VIRUSES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDermott, Will & Emery
; STREET: 600 13th Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3096
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/194,613
; FILING DATE: 30-NOV-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Bucca Ph.D., Daniel
; REGISTRATION NUMBER: 42,368
; REFERENCE/DOCKET NUMBER: 50179-061
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-756-8000
; TELEFAX: 202-756-8087
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6534 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-194-613-4

Query Match 2.4%; Score 60; DB 3; Length 6534;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 661 TTGACCCAGCGGTCTACAGAGTCTGCGCGCGGTTCGAGGACTCGAAGATCCCTTG 720
Db TTGACCCAGCGGTCTACAGAGTCTGCGCGCGGTTCGAGGACTCGAAGATCCCTTG 719

RESULT 10
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 2.0%; Score 49; DB 3; Length 4403765;
Best Local Similarity 47.3%; Pred. No. 0.11;
Matches 148; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 1251 AGCGCGCTGCACTCGCGTTTGTGTGGCAGAGTTTGCCACAGGTTGGCAGCTCGCGCTGCA 1310
DB 1213114 ATCGCGCGGAGCGCCAGCGCGCGCGCGGTGGGGCGGAGGACGAGTGGGGCGGA 1213173

QY 1311 GGCACCTCCAGCTGGGAGCAGGATCTCGGTGGCTACTCCTGGGTGGCTACTCCTGGGCGCCACACGGT 1370
DB 1213174 GCGAACGGTGGGTGGTTGTTGGCAATGGTGGTGGCGGGCTCCGGTGGCGACGCGCGC 1213233

QY 1371 ACTACTTCCCAGTGGCTCGGTAGCTAGCTTCTCCCTGAGGTTTCGCCCTTGAGCGC 1430
DB 1213234 GCGCGCGCGGTGGCGGTCCCGCAACTTGGGCTCGCGGGGGTATCAACGCCCGCGCC 1213293

QY 1431 TAGACCCGAAACAGCGCTCTTGGACCGACTTCGCTTCGCGAGGAGACACCGTCACTTTC 1490
DB 1213294 GGTAACCCCGGACGCGCTCGGTGGCATCGCGGTGGCGGTGGCGCGGACCGCC 1213353

QY 1491 CGCGAGTGGCGGTGGAGAGTGGTGTGTACCAACACCCCGCGCGCGCGCGCGCGC 1550
DB 1213354 GGGCTGTTCGGGACGGTGGGGTGGTGGCGCGGTGGTGGCGCGCGCGCGCGCTTC 1213413

QY 1551 CCCACCTTACCG 1563
DB 1213414 GCGGCATCAGCG 1213426

RESULT 11
US-09-103-840A-1
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 2.0%; Score 49; DB 3; Length 4411529;
Best Local Similarity 47.3%; Pred. No. 0.11;
Matches 148; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 1251 AGCGCGCTGCACTCGCGTTTGTGTGGCAGAGTTTGCCACAGGTTGGCAGCTCGCGCTGCA 1310
DB 1212977 ATCGCGCGGAGCGCGCGCGCGCGGTGGGGCGGAGGACGAGTGGGGCGCGGA 1213036

QY 1311 GGCACCTCCAGCTGGGAGCAGGATCTCGGTGGCTACTCCTGCGGCGCCACACGGT 1370
DB 1213037 GCGAACGGTGGGTGGTTGTTGGCAATGGTGGTGGCGGGCTCCGGTGGCGACGCGCGC 1213096

QY 1371 ACTACTTCCCAGTGGCTCGGTAGCTAGCTTCTCCCTGAGGTTTCGCCCTTGAGCGC 1430
DB 1213097 GCGCGCGCGGTGGCGGTGGCGCAACTTGGGCTCGGCGGGGTATCAACGCCCGCGCC 1213156

QY 1431 TAGACCCGAAACAGCGCTCTTGGACCGACTTCGCTTCGCGAGGAGACACCGTCACTTTC 1490
DB 1213157 GGTAACCCCGGACGCGCTCGGTGGCATCGCGGTGGCGGTGGCGGGGACCGCC 1213216

QY 1491 CGCGAGTGGCGGTGGAGAGTGGTGTGTACCAACACCCCGCGCGCGCGCGCGCGC 1550
DB 1213217 GGGCTGTTCGGGACGGTGGGGCTGGTGGGCGCGGTGGTGGCGCGCGCGCGCGCTTC 1213276

QY 1551 CCCACCTTACCG 1563
DB 1213277 GCGGCATCAGCG 1213289

RESULT 12
US-09-252-991A-13089/c
Sequence 13089, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 13089
LENGTH: 2136
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13089

Query Match 1.9%; Score 47.8; DB 4; Length 2136;
Best Local Similarity 46.3%; Pred. No. 0.022;
Matches 157; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 1766 CGTTTTCCAGCTCACGCCAGCCAGCTCCTTTGGCGCGCTTCCTTTCAACAATCCGGGTTA 1825
DB 1381 CCTGCTGCGCACCTTGAAGCCCTCGGCTTCGCGTGGTGGCTAGCGCTGACGCTTTG 1322

QY 1826 TGAGCGCACAGGACCTCCCGGACTACACTGGCATCCGTGACTCATTCGACGAGACAT 1885
DB 1321 CAACGGCATGTCCGGCGCCCTCGACCCGCGATCCAGCAGGAGATCGTCGAGCGGACCT 1262

QY 1886 GTCCACCGCTGGCGCCACTTCGCTCACTCTCCCACTCCTCGAGTATCGTCACTAAGAC 1945
DB 1261 GTACGCCACCGCGGTGCTCTCGGCAACCGCACTTCGACGGGCGCATCCACCCCTACG 1202

QY 1946 CTACCGAGGTTGGGAAGCGGTCAACAGCGCTTTCGCGCAATTCGCGCACGC 2005
DB 1201 CAAGCAGGCTTCTCGCTCGCTCGCGCGCTGGTGGTGGCTACGCCATCGCGGGACCAT 1142

QY 2006 GGGCTCTCTCAAGATGAGGAGATCTCTGCTTCGCGGACGACCTGGCCACCGCTCAC 2065
DB 1141 CGGCTTCGACATCGAGCGGACGTGCTCGGCGGTGGTGGACGGAAGGAGATCGCGCTGAA 1082

QY 2065 AGGTGTCTACCGCGCCACTGACAACTTCGCGCGCGCGCT 2104
DB 1081 GGACCTCTGCGCAGCGACGAGGATGACGCGGTGGT 1043

RESULT 13
US-09-252-991A-12703
; Sequence 12703, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12703
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12703

Query Match 1.9%; Score 47.8; DB 4; Length 2406;
Best Local Similarity 46.3%; Pred. No. 0.023;
Matches 157; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 1766 CGTTTCCAGCTCACGCCAGCCAGCTCTCTTTGGCGCGTTTCTTCAACAATCCGGTTA 1825
Db 1138 CTGCTGCGGCACCTGGAGCCCTCGGCTTGGCGTGGCTAGCCTGTACGTCTTG 1197

QY 1826 TGAGGCGCACGCGACTCCCGGACTACATCGTCATCTCGTACTATTCGACAGAACAT 1885
Db 1198 CAACGCGCATGTCGGCGCCCTCGACCCGCGATCCAGCAGGAGATCGTCGAGCGGACCT 1257

QY 1886 GTCCACCGTGTGGCCACTTCGCTCACTCTCCCACTCTCGAGTATCGTCACTAAGAC 1945
Db 1258 GTACGCGCACCGGCTGCTCTCGGCAACCGCACTTCGACGGCGCATCCACCCCTACGC 1317

QY 1946 CTACAGGGTGTGGAGGCGTCACGAACGTCAACAGCCTTTTCGGGCAATTCGCGCACGC 2005
Db 1318 CAAGAGGCTTCTCGCTCGCCGCGCTGGTGGTGGCTAGCCATCGCGGACCAT 1377

QY 2006 GGGCTCTCTCAAGATGAGAGATCTCTGCTCGCCGCGAGCCTTGGCCACCCGCTCTAC 2065
Db 1378 CGCTTCGACATCGAGCGGACGCTGCTCGGCGTGGTGGACGCGAGGATCGGCTGAA 1437

QY 2066 AGTGTCTACCCCGCACTGACAACTTCGCGCGCGCT 2104
Db 1438 GGACCTCTGGCGAGCGAGGAGATCGACGCGGTGGT 1476

RESULT 14
US-09-252-991A-12851
; Sequence 12851, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12851
; LENGTH: 2616
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12851

Query Match 1.9%; Score 47.8; DB 4; Length 2616;
Best Local Similarity 46.3%; Pred. No. 0.023;
Matches 157; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 1766 CGTTTCCAGCTCACGCCAGCCAGCTCTCTTTGGCGCGTTTCTTCAACAATCCGGTTA 1825
Db 1389 CTGCTGCGGCACCTGGAGCCCTCGGCTTGGCGTGGCTAGCCTGTACGTCTTG 1448

QY 1826 TGAGGCGCACGCGACTCCCGGACTACATCGTCATCTCGTACTATTCGACAGAACAT 1885
Db 1449 CAACGCGCATGTCGGCGCCCTCGACCCGCGATCCAGCAGGAGATCGTCGAGCGGACCT 1508

QY 1886 GTCCACCGTGTGGCCACTTCGCTCACTCTCCCACTCTCGTACTATTCGCTAAGAC 1945
Db 1509 GTACGCGCACCGGCTGCTCTCGGCAACCGCACTTCGACGGCGCATCCACCCCTACGC 1568

QY 1946 CTACAGGGTGTGGAGGCGTCACGAACGTCAACAGCCTTTTCGGGCAATTCGCGCACGC 2005
Db 1569 CAAGAGGCTTCTCGCTCGCCGCGCTGGTGGTGGCTAGCCATCGCGGACCAT 1628

QY 2006 GGGCTCTCTCAAGATGAGGAGATCTCTGCTCGCCGCGAGCCTTGGCCACCCGCTCTAC 2065
Db 1629 CGCTTCGACATCGAGCGCGAGCTGCTCGGCGTGGTGGACGCGAGGATCGGCTGAA 1688

QY 2066 AGTGTCTACCCCGCACTGACAACTTCGCGCGCGCT 2104
Db 1689 GGACCTCTGGCGAGCGAGGAGATCGACGCGGTGGT 1727

RESULT 15
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; Zip: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pr29pt-fls
US-08-232-463-14

Query Match      1.9%; Score 46.6; DB 1; Length 7218;
Best Local Similarity 1.8%; Pred. No. 0.062;
Matches 7; Conservative 219; Mismatches 153; Indels 0; Gaps 0;

Qy 1768 TTTTCAGCTCAGCGCAGCTCCTTTGGCGCGTTTCCTTCAACAAATCCGGGTATG 1827
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1067 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1126
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1828 ACGCACACGCGACCTCCGGACTACACTGGCATCCGTGACTCATTGACCAAGACATGT 1887
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1127 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1186
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1888 CCACGCTGTGGCCCACTTCCGCTCACTCTCCACTCTCGCAGTATCGTCACTAAGACCT 1947
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1187 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1246
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1948 ACCAGGTTGGAAGCGTCAGACGTCAACACGCTTTTCGGCCAAATTCGGCGACGCGG 2007
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1247 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1306
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 2008 GCCTCTCAAGATGAGGAGATCCTCTGCTCGCGACGACCTGGCCACCCGCTCTCACAG 2067
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1307 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1366
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 2068 GTGCTACCCGCGCACTGACAACTTCGCGCGCGCGTTTCTGCTTCGCGCGCAACATGC 2127
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1367 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1426
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 2128 TGTCTCCGTGCTGAAGTC 2146
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1427 YYYYYYGTACCAAAATTC 1445
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: January 12, 2004, 22:35:29
Job time : 184 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 21:07:11 ; Search time 823 Seconds
(without alignments)
10455.814 Million cell updates/sec

Title: US-09-991-262-47

Perfect score: 2478

Sequence: 1 GTTTTCTTCTTTACCAAG.....CTGCGTCCCTCGGATACCA 2478

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2276164 seqs, 1736306516 residues

Total number of hits satisfying chosen parameters: 4552328

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2478	100.0	2478	11	US-09-991-262-47
2	2478	100.0	2478	11	US-09-991-262-47
3	2467	99.6	2479	11	US-09-991-262-49
4	71.2	2.9	5312	11	US-09-991-262-39
5	71.2	2.9	5312	11	US-09-991-262-41
6	71.2	2.9	5312	11	US-09-991-262-43
7	71.2	2.9	5368	11	US-09-991-262-45
8	48.4	2.0	750	15	US-10-156-761-6810
9	48.4	2.0	9025608	15	US-10-156-761-1
10	48.2	1.9	1416	15	US-10-156-761-6911
11	47.8	1.9	1926	15	US-10-156-761-2231
12	47.8	1.9	9025608	15	US-10-156-761-1
13	3570	1.9	3570	15	US-10-156-761-6645
14	46.4	1.9	2679	15	US-10-156-761-176
15	46.2	1.9	480	15	US-10-156-761-5243

16	46.2	1.9	1194	15	US-10-156-761-2770
17	45.6	1.8	1079	15	US-10-156-761-3854
18	45.4	1.8	1717	15	US-10-156-761-1387
19	45.2	1.8	1323	12	US-10-369-493-42288
20	44.8	1.8	3303	15	US-10-156-761-5384
21	44.6	1.8	1044	15	US-10-156-761-539
22	44.6	1.8	3285	13	US-10-228-063-46
23	44.2	1.8	657	15	US-10-156-761-7488
24	44.2	1.8	1314	9	US-09-815-242-7711
25	44.2	1.8	1211	15	US-10-156-761-6420
26	44	1.8	1176	15	US-10-156-761-3439
27	44	1.8	1971	12	US-10-369-493-42357
28	43.8	1.8	815	13	US-10-084-843-139
29	43.8	1.8	815	13	US-10-193-002-134
30	43.8	1.8	831	15	US-10-156-761-1308
31	43.6	1.8	1242	15	US-10-156-761-6018
32	43.6	1.8	1371	15	US-10-156-761-6160
33	43.4	1.8	774	13	US-10-203-708-19
34	43.4	1.8	774	13	US-10-029-386-20965
35	43.4	1.8	1371	15	US-10-156-761-5235
36	43.2	1.7	520	15	US-10-184-644-332
37	43.2	1.7	520	15	US-10-184-644-332
38	43.2	1.7	711	12	US-10-369-493-42329
39	43	1.7	3705	12	US-10-369-493-42324
40	42.8	1.7	1141	15	US-10-184-644-120
41	42.8	1.7	1141	15	US-10-184-644-120
42	42.6	1.7	1296	15	US-10-156-761-3361
43	42.6	1.7	2105	13	US-10-289-757-125
44	42.4	1.7	691	13	US-10-140-472-16
45	42.4	1.7	691	13	US-10-141-761-16

ALIGNMENTS

RESULT 1

US-09-991-262-47
; Sequence 47, Application US/09991262
; Publication No. US20030041349A1
; GENERAL INFORMATION:
; APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and Their Uses in
; Protecting Plants
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/991,262
; FILING DATE: 20-No. US20030041349A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,238
; FILING DATE: 20-JAN-1999
; APPLICATION NUMBER: US 08/485,355
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/440,522
; FILING DATE: 12-MAY-1995
; APPLICATION NUMBER: US 08/089,372
; FILING DATE: 08-JUL-1993
; APPLICATION NUMBER: AU PL4081/92
; FILING DATE: 14-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.

REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-586311-3/RFT/DSS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2478 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 283..753
SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-991-262-47
Query Match 100.0%; Score 2478; DB 11; Length 2478;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTTTTCTTTTACCAAGTGTGTAATAATTTAAACAAAGAGAAACCCAGGACCGTAA 60
DB 1 GTTTTCTTTTACCAAGTGTGTAATAATTTAAACAAAGAGAAACCCAGGACCGTAA 60
QY 61 CCGGCCCCATACACCTCGAGTCCGTGACACCGGATTATAGTCGCCCAACACACGGC 120
DB 61 CCGGCCCCATACACCTCGAGTCCGTGACACCGGATTATAGTCGCCCAACACACGGC 120
QY 121 GCGTTTCCGACACTCTCGAGAGTCTGTTGGAGTTTCGTCGTGACACCGGTTGGCA 180
DB 121 GCGTTTCCGACACTCTCGAGAGTCTGTTGGAGTTTCGTCGTGACACCGGTTGGCA 180
QY 181 GTCGACAGACGCTTCGGGACCACTAGAACCTCTCGAGGCGGCACACACAGCACACA 240
DB 181 GTCGACAGACGCTTCGGGACCACTAGAACCTCTCGAGGCGGCACACACAGCACACA 240
QY 241 CCGCCTTAGCTGACCTACCGGAGCGTTGATAGCGCGATTATAGCGGAGCACACCATC 300
DB 241 CCGCCTTAGCTGACCTACCGGAGCGTTGATAGCGCGATTATAGCGGAGCACACCATC 300
QY 301 GCGCACTCATCATATTACACCGGTTACACCTTGCCTTAATACCCCTGAACTGAA 360
DB 301 GCGCACTCATCATATTACACCGGTTACACCTTGCCTTAATACCCCTGAACTGAA 360
QY 361 GCAGGATGGAGATGCTGGAGTGGCTGACAGGACCTCACACCGTCCGGAAACCGGTA 420
DB 361 GCAGGATGGAGATGCTGGAGTGGCTGACAGGACCTCACACCGTCCGGAAACCGGTA 420
QY 421 AGCTTCGGGTGACGCGCAACACCGTCAACCGTCAATGTTAGAGAAACCAACCGGTCGA 480
DB 421 AGCTTCGGGTGACGCGCAACACCGTCAACCGTCAATGTTAGAGAAACCAACCGGTCGA 480
QY 481 CCGGAAGGCAAGTTTCTCCCTGACAAATTCACCGCTGCTGCAAGAACCTCGGCCAA 540
DB 481 CCGGAAGGCAAGTTTCTCCCTGACAAATTCACCGCTGCTGCAAGAACCTCGGCCAA 540
QY 541 GCTTCGACGCAACACCGTCACTTTCCTCCGCTAACATCTAGCATGCCGAATCCGGA 600
DB 541 GCTTCGACGCAACACCGTCACTTTCCTCCGCTAACATCTAGCATGCCGAATCCGGA 600
QY 601 ATTGGGCCAAGGAAAGATCGACTCGACTCCGATTCATCCGCTGGTACTTCAAGTACC 660
DB 601 ATTGGGCCAAGGAAAGATCGACTCGACTCCGATTCATCCGCTGGTACTTCAAGTACC 660
QY 661 TTGACCCAGGGGTGCTACAGAGTCTGCGGGCGCGTCCGCGAGTACTCGAAGATCCCTG 720
DB 661 TTGACCCAGGGGTGCTACAGAGTCTGCGGGCGCGTCCGCGAGTACTCGAAGATCCCTG 720
QY 721 ACGGCTCTGTAAGTTCTCGCTCGACGAGAGATAAGAGATCTATAACGAGGAGTGC 780
DB 721 ACGGCTCTGTAAGTTCTCGCTCGACGAGAGATAAGAGATCTATAACGAGGAGTGC 780

DB 721 ACGGCTCTGTAAGTTCTCGCTCGACGAGAGATAAGAGATCTATAACGAGGAGTGC 780
QY 781 CCGTCGTCACTGACGTGTCCGTCCTCGACGCGCGCAGTGGAGCTCTCGATTTTCT 840
DB 781 CCGTCGTCACTGACGTGTCCGTCCTCGACGCGCGCAGTGGAGCTCTCGATTTTCT 840
QY 841 CTTTCCGATGTTTCAAGAACCGCTACGTCCGCTAGCGAAACGTCGAGAACGAGAGTGT 900
DB 841 CTTTCCGATGTTTCAAGAACCGCTACGTCCGCTAGCGAAACGTCGAGAACGAGAGTGT 900
QY 901 CCGTCGAGTGTGTAACGACCTCATCGAGTGGCTCAACAATCTCGCCGACGCTGATG 960
DB 901 CCGTCGAGTGTGTAACGACCTCATCGAGTGGCTCAACAATCTCGCCGACGCTGATG 960
QY 961 TCGTTGACTCTGAACAGTGGATTAACTTCAACAATGACACACGACTACTACGTCGCCATCC 1020
DB 961 TCGTTGACTCTGAACAGTGGATTAACTTCAACAATGACACACGACTACTACGTCGCCATCC 1020
QY 1021 GCGTTCTACGTCCAACTTACGACGTTCCAGACCCCAAGAGGCGCTTGTTCGACAGTCT 1080
DB 1021 GCGTTCTACGTCCAACTTACGACGTTCCAGACCCCAAGAGGCGCTTGTTCGACAGTCT 1080
QY 1081 CAGACTACCGCTCTCACTTATAAGGCGATAACATGTGAAGCCCAACATGCCACACTCGTCG 1140
DB 1081 CAGACTACCGCTCTCACTTATAAGGCGATAACATGTGAAGCCCAACATGCCACACTCGTCG 1140
QY 1141 ACCAAGGCTTTGGATCGGCGCCAGTACGCTCTCACCCGACTAGCTTACCGCAGTACG 1200
DB 1141 ACCAAGGCTTTGGATCGGCGCCAGTACGCTCTCACCCGACTAGCTTACCGCAGTACG 1200
QY 1201 AGCTCAGCGAGCGCTTACGCTCTGACACATTTGACCTTCGCCAGACCATCCAGCGCGCTG 1260
DB 1201 AGCTCAGCGAGCGCTTACGCTCTGACACATTTGACCTTCGCCAGACCATCCAGCGCGCTG 1260
QY 1261 CACTCGCGTTTGTGGGAGGTTTGGCCAGGCTGCGCTGCGGCTGAGGCACTCCAG 1320
DB 1261 CACTCGCGTTTGTGGGAGGTTTGGCCAGGCTGCGCTGCGGCTGAGGCACTCCAG 1320
QY 1321 CCGTGGAGCGAGCATCTCGGCTGCTACCTCAGCTGGCGCCACAAACGCTACTCTTCC 1380
DB 1321 CCGTGGAGCGAGCATCTCGGCTGCTACCTCAGCTGGCGCCACAAACGCTACTCTTCC 1380
QY 1381 CAGCTGGCTTCGTTAGCTACGTTCTCCCTGAGGGTTTCGCCCTTACGAGCGCTACGACCGGA 1440
DB 1381 CAGCTGGCTTCGTTAGCTACGTTCTCCCTGAGGGTTTCGCCCTTACGAGCGCTACGACCGGA 1440
QY 1441 ACGACGGCTTCGGACCGACTTCGCTTCGCCAGGAGACACCGTCACTTTCCGCGAGGTCG 1500
DB 1441 ACGACGGCTTCGGACCGACTTCGCTTCGCCAGGAGACACCGTCACTTTCCGCGAGGTCG 1500
QY 1501 CCGTCGACGAGTTCGTTGACCAACCGCGCGGCGGAGCGCCGCCACCTTCA 1560
DB 1501 CCGTCGACGAGTTCGTTGACCAACCGCGCGGCGGAGCGCCGCCACCTTCA 1560
QY 1561 CCGTCGAGAGTTCGCTTCAACCGCTTACCAACCGCTTACCAACCGCTTACCAACCGCTTAC 1620
DB 1561 CCGTCGAGAGTTCGCTTCAACCGCTTACCAACCGCTTACCAACCGCTTACCAACCGCTTAC 1620
QY 1621 AGACTCGACCTTCCTCGTAGGCTCGAACTCCCTATGCGACCTGCTGACTTTGGACAGA 1680
DB 1621 AGACTCGACCTTCCTCGTAGGCTCGAACTCCCTATGCGACCTGCTGACTTTGGACAGA 1680
QY 1681 CCGTCGCAACCAACCGGAGATCGAGCAGTCTTAAAGAAACACTTTGGCTGCTATT 1740
DB 1681 CCGTCGCAACCAACCGGAGATCGAGCAGTCTTAAAGAAACACTTTGGCTGCTATT 1740
QY 1741 TGGTCCACTTCAAAATGCGAAACCCGCTTTCAGCTCACGCCAGCGAGCTCTTTGGCG 1800
DB 1741 TGGTCCACTTCAAAATGCGAAACCCGCTTTCAGCTCACGCCAGCGAGCTCTTTGGCG 1800
QY 1801 CCGTTTCTTCAACAAATCCGGGTTATAGCGGACACCGGACCTCCGGGACTACCTGGCA 1860
DB 1801 CCGTTTCTTCAACAAATCCGGGTTATAGCGGACACCGGACCTCCGGGACTACCTGGCA 1860

1861 TCCTGACTCATTCGACCGAAGACATGTCCACCGCTGTGCGCCCACTTCCTCCCTCACTCTCCC 1920
1861 TCCTGACTCATTCGACCGAAGACATGTCCACCGCTGTGCGCCCACTTCCTCCCTCACTCTCCC 1920
1921 ACTCCTGAGTATCGTCACTAGACCTACAGGTTGGGAAGCGCTCAGAACGTCACAA 1980
1921 ACTCCTGAGTATCGTCACTAGACCTACAGGTTGGGAAGCGCTCAGAACGTCACAA 1980
1981 CGCTTTTCGGCCAAATTCGGCGCAGCGGGCTCTCTCAAGATGAGGAGATCTCTGCTCG 2040
1981 CGCTTTTCGGCCAAATTCGGCGCAGCGGGCTCTCTCAAGATGAGGAGATCTCTGCTCG 2040
2041 CGGAGACCTGCGCCACCTGCTCAGAGTGTCTACCCCGCCACTGACAACTTCGGGGCG 2100
2041 CGGAGACCTGCGCCACCTGCTCAGAGTGTCTACCCCGCCACTGACAACTTCGGGGCG 2100
2101 CGCTTTTCGGCCAAATTCGGCGCAGCGGGCTCTCTCAAGATGAGGAGATCTCTGCTCG 2160
2101 CGCTTTTCGGCCAAATTCGGCGCAGCGGGCTCTCTCAAGATGAGGAGATCTCTGCTCG 2160
2161 CCATCATCAAGTCCGTTGGCGGAGACTGCGCTCGGCGGCTCAGTCCGGCTCGGAGC 2220
2161 CCATCATCAAGTCCGTTGGCGGAGACTGCGCTCGGCGGCTCAGTCCGGCTCGGAGC 2220
2221 TACCGGACTGTATGAGTGTACAGGAGAGATTCGCGCGCTGTCCGCGCGCCGAG 2280
2221 TACCGGACTGTATGAGTGTACAGGAGAGATTCGCGCGCTGTCCGCGCGCCGAG 2280
2281 CGGCGCGCGCGCGCTGTCCGCGGAGATTCGCTCGCTCTCTGCTCGCTCTGCTCGCTAA 2340
2281 CGGCGCGCGCGCGCTGTCCGCGGAGATTCGCTCGCTCTCTGCTCGCTCTGCTCGCTAA 2340
2341 ACGGCTGTGTCGCGACATTCAGCGTACCTTAAAGACTCTGTGAGTCCCGCTCGTTACA 2400
2341 ACGGCTGTGTCGCGACATTCAGCGTACCTTAAAGACTCTGTGAGTCCCGCTCGTTACA 2400
2401 CGACGGTGTGCGCGGTTGATTCATTCATTCGCGGCAAGGAGGAGGAGTGTAGTCTCT 2460
2401 CGACGGTGTGCGCGGTTGATTCATTCATTCGCGGCAAGGAGGAGGAGTGTAGTCTCT 2460
2461 GGTGCTCTCGGATACCA 2478
2461 GGTGCTCTCGGATACCA 2478

RESULT 2

US-09-991-262-49
; Sequence 49, Application US/09991262-
; Publication No. US20030041349A1
; GENERAL INFORMATION:
; APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and Their Uses in
; Protecting Plants
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Flehr Hohbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/991,262
; FILING DATE: 20-No. US20030041349A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,238

FILING DATE: 20-JAN-1999
APPLICATION NUMBER: US 08/485,355
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/440,522
FILING DATE: 12-MAY-1995
APPLICATION NUMBER: US 08/089,372
FILING DATE: 08-JUL-1993
APPLICATION NUMBER: AU PL4081/92
FILING DATE: 14-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-58631-3/RFT/DSS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2478 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 366..2306
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-09-991-262-49
Query Match 100.0%; Score 2478; DB 11; Length 2478;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTTTTCTTTCTTTTACCAAGTGTGTAATAATTTAAACAAGAAAGAAACCGGACCTTAA 60
DB 1 GTTTTCTTTCTTTTACCAAGTGTGTAATAATTTAAACAAGAAAGAAACCGGACCTTAA 60
QY 61 CCGCGCCCTTACACACCTTCGAGTCCGTGACCAACCGGATTATACGTCGCCACCAACCGC 120
DB 61 CCGCGCCCTTACACACCTTCGAGTCCGTGACCAACCGGATTATACGTCGCCACCAACCGC 120
QY 121 GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTCCTCGTGACCAACCGGTTGGCA 180
DB 121 GCCTTTTCCGACCACTCTCGAGAGTCTGTTGGAGTTCCTCGTGACCAACCGGTTGGCA 180
QY 181 GTGACAGACCGTTCGGACCACTAGAACCTTCCTCGAGCGGACGACACACAGCACACA 240
DB 181 GTGACAGACCGTTCGGACCACTAGAACCTTCCTCGAGCGGACGACACACAGCACACA 240
QY 241 CCGCTTAGCTGCACCTACGCGAGCGTGTGATAGCGCGATTATGAGCGGAGCACACCATC 300
DB 241 CCGCTTAGCTGCACCTACGCGAGCGTGTGATAGCGCGATTATGAGCGGAGCACACCATC 300
QY 301 GCCCACTCCATCACATTACCAACCGGTTACACCTTTCCTTAATACCCCTGAACTGAA 360
DB 301 GCCCACTCCATCACATTACCAACCGGTTACACCTTTCCTTAATACCCCTGAACTGAA 360
QY 361 CGAGATGGGAGATGCTGGAGTGGGTGACAGGACCTTCACACCGTTCGGGAACCGTA 420
DB 361 CGAGATGGGAGATGCTGGAGTGGGTGACAGGACCTTCACACCGTTCGGGAACCGTA 420
QY 421 ACGTTCGGGTTCAGCGCAACACCGTTCACCGTCAATGTTAGAGAAACCAACCGGTCGGA 480
DB 421 ACGTTCGGGTTCAGCGCAACACCGTTCACCGTCAATGTTAGAGAAACCAACCGGTCGGA 480
QY 481 CCGGAAGCAAGTTTCTCCCTGACATTTTCCCGTGTGTCACAGACCTTCGGGAAA 540
DB 481 CCGGAAGCAAGTTTCTCCCTGACATTTTCCCGTGTGTCACAGACCTTCGGGAAA 540
QY 541 GCCTTGACGCAACACCGTCTCTTTCCCGCTTACATCTTAGCATGCCGGAATTCGGA 600
DB 541 GCCTTGACGCAACACCGTCTCTTTCCCGCTTACATCTTAGCATGCCGGAATTCGGA 600

601 ATTTGGCCAAAGGAAGATCGACCTCGATCCGATTCCATCGGTGGTACTTCAAGTACC 660
Db |
601 ATTTGGCCAAAGGAAGATCGACCTCGATCCGATTCCATCGGTGGTACTTCAAGTACC 660
Qy |
661 TTGACCCAGCGGTGCTACAGAGTCTGCGCGCGCGTGGCGAGTACTCGAAGATCCCTG 720
Db |
661 TTGACCCAGCGGTGCTACAGAGTCTGCGCGCGCGTGGCGAGTACTCGAAGATCCCTG 720
Qy |
721 ACGGCTCTGTAAGTTCTCGTTCGACGAGAGATAGAGAGATCTATACAGAGATGCC 780
Db |
721 ACGGCTCTGTAAGTTCTCGTTCGACGAGAGATAGAGAGATCTATACAGAGATGCC 780
Qy |
781 CGGTGCTACTGACGTGCTCGTTCGACGAGAGATAGAGAGATCTATACAGAGATGCC 840
Db |
781 CGGTGCTACTGACGTGCTCGTTCGACGAGAGATAGAGAGATCTATACAGAGATGCC 840
Qy |
841 CTTTTCCGATGTTCAAGACCGCTACGTTCGCGGTAGCGAACTCGAGAAACAGAGATGT 900
Db |
841 CTTTTCCGATGTTCAAGACCGCTACGTTCGCGGTAGCGAACTCGAGAAACAGAGATGT 900
Qy |
901 CGCTGACCTGTTCAACGACCTCATCGAGTGGCTCAACAACTCTGCGCGACTGGCGTTATG 960
Db |
901 CGCTGACCTGTTCAACGACCTCATCGAGTGGCTCAACAACTCTGCGCGACTGGCGTTATG 960
Qy |
961 TCGTTGACTCTGAACAGTGGATTAACCTTCAACAACTGACACCACTGCTACGTCGGCATCC 1020
Db |
961 TCGTTGACTCTGAACAGTGGATTAACCTTCAACAACTGACACCACTGCTACGTCGGCATCC 1020
Qy |
1021 GGTTCCTACGTCACACTACGAGTTCAGAGTTCAGACCCACAGAGGCGCTTGTTCGACAGTCT 1080
Db |
1021 GGTTCCTACGTCACACTACGAGTTCAGAGTTCAGACCCACAGAGGCGCTTGTTCGACAGTCT 1080
Qy |
1081 CAGACTACCGCTCTACTTAAAGGCGATAACATGTGAAGCCAAACATGCCAAACACTCGTGG 1140
Db |
1081 CAGACTACCGCTCTACTTAAAGGCGATAACATGTGAAGCCAAACATGCCAAACACTCGTGG 1140
Qy |
1141 ACCAAGGCTTTGGATCGGCGGCGAGTACGCTCTACCCCGACTAGCCTACCGCAGTACG 1200
Db |
1141 ACCAAGGCTTTGGATCGGCGGCGAGTACGCTCTACCCCGACTAGCCTACCGCAGTACG 1200
Qy |
1201 ACCTCAGCAGGCTACGCTCTGACACTTTGACCTTCGACAGACCATCAGCGCGCTG 1260
Db |
1201 ACCTCAGCAGGCTACGCTCTGACACTTTGACCTTCGACAGACCATCAGCGCGCTG 1260
Qy |
1261 CACTCGCTTTGTGTTGGCAGGTTTGGCAGGTTGGCAGGTTGGCAGGTTGGCAGGTTGGCAG 1320
Db |
1261 CACTCGCTTTGTGTTGGCAGGTTTGGCAGGTTGGCAGGTTGGCAGGTTGGCAGGTTGGCAG 1320
Qy |
1321 CTTGGAGCAGGATCCTCGGTGGTACTCTACCTGGCGGCGACACGCTACTTTTCC 1380
Db |
1321 CTTGGAGCAGGATCCTCGGTGGTACTCTACCTGGCGGCGACACGCTACTTTTCC 1380
Qy |
1381 CAGCTGGCTCCGTTAGTACTGTTCTCCCTGAGGTTTGGCGGCTAGCGCTACGACCCGA 1440
Db |
1381 CAGCTGGCTCCGTTAGTACTGTTCTCCCTGAGGTTTGGCGGCTAGCGCTACGACCCGA 1440
Qy |
1441 ACGACGGCTCTTGGACCGACTTCGTTTCGCGAGGACACCGTCACTTTTCGCGAGGTGG 1500
Db |
1441 ACGACGGCTCTTGGACCGACTTCGTTTCGCGAGGACACCGTCACTTTTCGCGAGGTGG 1500
Qy |
1501 CCGTCAGCAGGTCGTTGTGACCAACACACCCCGCGCGCGCGAGCGGCGGCGGCGGCGGCG 1560
Db |
1501 CCGTCAGCAGGTCGTTGTGACCAACACACCCCGCGCGCGCGAGCGGCGGCGGCGGCGGCG 1560
Qy |
1561 CCGTGAGGTGCGGCTTCAACGCTTACACCAACACCGTGTAGGACACGCTCTTAG 1620
Db |
1561 CCGTGAGGTGCGGCTTCAACGCTTACACCAACACCGTGTAGGACACGCTCTTAG 1620
Qy |
1621 AGACTCGACCTCTCTCGTAGGCTCGAATCCCTATGCCACCTGCTGCTTTGGACAGA 1680
Db |
1621 AGACTCGACCTCTCTCTCGTAGGCTCGAATCCCTATGCCACCTGCTGCTTTGGACAGA 1680

Qy |
1681 CGGTTCCTTCAACAACTTCGACGAGATCGAGAGTTCGCTTTTAAAGAAACACTTGGCTCTATT 1740
Db |
1681 CGGTTCCTTCAACAACTTCGACGAGATCGAGAGTTCGCTTTTAAAGAAACACTTGGCTCTATT 1740
Qy |
1741 TGGTTCCTTCAACAACTTCGACGAGATCGAGAGTTCGCTTTTAAAGAAACACTTGGCTCTATT 1800
Db |
1741 TGGTTCCTTCAACAACTTCGACGAGATCGAGAGTTCGCTTTTAAAGAAACACTTGGCTCTATT 1800
Qy |
1801 CCGTTCCTTCAACAACTTCGACGAGATCGAGAGTTCGCTTTTAAAGAAACACTTGGCTCTATT 1860
Db |
1801 CCGTTCCTTCAACAACTTCGACGAGATCGAGAGTTCGCTTTTAAAGAAACACTTGGCTCTATT 1860
Qy |
1861 TCGGTGACTCATTCGACGAGATCGAGAGTTCGCTTTTAAAGAAACACTTGGCTCTATT 1920
Db |
1861 TCGGTGACTCATTCGACGAGATCGAGAGTTCGCTTTTAAAGAAACACTTGGCTCTATT 1920
Qy |
1921 ACTCTGAGTATCGTCACTAAGACCTACAGGTTGGGAAGGCTCAGACGTCACACA 1980
Db |
1921 ACTCTGAGTATCGTCACTAAGACCTACAGGTTGGGAAGGCTCAGACGTCACACA 1980
Qy |
1981 CGCTTTTCGCGCAATTCGCGCAGCGGCTCTCAAGAAATGAGAGATCTCTTCCCTCG 2040
Db |
1981 CGCTTTTCGCGCAATTCGCGCAGCGGCTCTCAAGAAATGAGAGATCTCTTCCCTCG 2040
Qy |
2041 CCGAGACTTCGCGCAATTCGCGCAGCGGCTCTCAAGAAATGAGAGATCTCTTCCCTCG 2100
Db |
2041 CCGAGACTTCGCGCAATTCGCGCAGCGGCTCTCAAGAAATGAGAGATCTCTTCCCTCG 2100
Qy |
2101 CGGTTCCTTTCGCGCAATTCGCGCAGCGGCTCTCAAGAAATGAGAGATCTCTTCCCTCG 2160
Db |
2101 CGGTTCCTTTCGCGCAATTCGCGCAGCGGCTCTCAAGAAATGAGAGATCTCTTCCCTCG 2160
Qy |
2161 CCATCATCAAGTTCGCGCAATTCGCGCAGCGGCTCTCAAGAAATGAGAGATCTCTTCCCTCG 2220
Db |
2161 CCATCATCAAGTTCGCGCAATTCGCGCAGCGGCTCTCAAGAAATGAGAGATCTCTTCCCTCG 2220
Qy |
2221 TACCCGAGTTCGCGCAATTCGCGCAGCGGCTCTCAAGAAATGAGAGATCTCTTCCCTCG 2280
Db |
2221 TACCCGAGTTCGCGCAATTCGCGCAGCGGCTCTCAAGAAATGAGAGATCTCTTCCCTCG 2280
Qy |
2281 CGCGCGCGCGCGCTCTCGTCCAAATTTAGTTTGTCTCGCTCTCTCTCTCTCTCTCTCTCT 2340
Db |
2281 CGCGCGCGCGCGCTCTCGTCCAAATTTAGTTTGTCTCGCTCTCTCTCTCTCTCTCTCTCT 2340
Qy |
2341 ACGCGGTGCTCCCGCACATTTAGTTTGTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
Db |
2341 ACGCGGTGCTCCCGCACATTTAGTTTGTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
Qy |
2401 CCGCGGTGCTCCCGCACATTTAGTTTGTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460
Db |
2401 CCGCGGTGCTCCCGCACATTTAGTTTGTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460
Qy |
2461 GCGTCCCTCGGATACCA 2478
Db |
2461 GCGTCCCTCGGATACCA 2478

RESULT 3

US-09-991-262-51
; Sequence 51, Application US/09991262
; Publication No. US20030041349A1
; GENERAL INFORMATION:
; APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and Their Uses in
; Protecting Plants
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/991,262
FILING DATE: 20-NO. US20030041349A1-2001
CLASSIFICATION DATA: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/234,238
FILING DATE: 20-JAN-1999
APPLICATION NUMBER: US 08/485,355
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/440,522
FILING DATE: 12-MAY-1995
APPLICATION NUMBER: US 08/089,372
FILING DATE: 08-JUL-1993
APPLICATION NUMBER: AU PL4081/92
FILING DATE: 14-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard P.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-58631-3/RFT/DSS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 2479 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 283..2307
SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-991-262-51

Query Match 99.6%; Score 2467; DB 11; Length 2479;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2478; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GTTTTCTTTTACCAAGTGTGGTAAATTTAAACAAAGAGAAACACGACCGTAA 60
DB 1 GTTTTCTTTTACCAAGTGTGGTAAATTTAAACAAAGAGAAACACGACCGTAA 60
QY 61 CCGGCGCTTACACACTCGAGTCCGTGACACCGGATTATACGTGCGCCACACACGGC 120
DB 61 CCGGCGCTTACACACTCGAGTCCGTGACACCGGATTATACGTGCGCCACACACGGC 120
QY 121 GCCTTTTCGACACTCTCGAGTGTGGTGGAGTTTCGTCGTCGACCAACCGTTGGCA 180
DB 121 GCCTTTTCGACACTCTCGAGTGTGGTGGAGTTTCGTCGTCGACCAACCGTTGGCA 180
QY 181 GTGACAGAGCTTCGGACCACTAGAACCTCTCGAGCGACGACACACACACACA 240
DB 181 GTGACAGAGCTTCGGACCACTAGAACCTCTCGAGCGACGACACACACACACA 240
QY 241 CCGCCTTAGCTGACCTACGAGCGGTTGATAGCGGATTATAGAGGAGCACCAATC 300
DB 241 CCGCCTTAGCTGACCTACGAGCGGTTGATAGCGGATTATAGAGGAGCACCAATC 300
QY 301 GCCCACTCCATCACATTACCAACCGGTTACACCTTGCCTTAATACCCCTGAACCTGAA 360
DB 301 GCCCACTCCATCACATTACCAACCGGTTACACCTTGCCTTAATACCCCTGAACCTGAA 360
QY 361 GCAGATGGAGATGCTGAGTGGGTGACAGGACCTCACAAACCGTTCGCGAAACCGGTA 420
DB 361 GCAGATGGAGATGCTGAGTGGGTGACAGGACCTCACAAACCGTTCGCGAAACCGGTA 420

QY 421 ACGTTCCGGTCAGCGCCAAACACCGTCAACCGTCAATAGGTAGAGAAACCAACCGCGTCGGA 480
DB 421 ACGTTCCGGTCAGCGCCAAACACCGTCAACCGTCAATAGGTAGAGAAACCAACCGCGTCGGA 480
QY 481 CGGGAAGCAAGTTTCTCCCTGACAAATTTACCGCTGCTGCACAAAGACCTTCGCGCAAA 540
DB 481 CGGGAAGCAAGTTTCTCCCTGACAAATTTACCGCTGCTGCACAAAGACCTTCGCGCAAA 540
QY 541 GCCTTGACGCCAAACACCGTCACTTT - CCGCGCTAAACATCTCTAGCATGCCGAAATTCGG 599
DB 541 GCCTTGACGCCAAACACCGTCACTTTCCCGCCGCTAACATCTCTAGCATGCCGAAATTCGG 600
QY 600 AATTGGGCAAGGAAAGATCGACTCGACTCCGATTCATCGCTGCTGATCTTCAAGTAC 659
DB 601 AATTGGGCAAGGAAAGATCGACTCGACTCCGATTCATCGCTGCTGATCTTCAAGTAC 660
QY 660 CTTGACCCAGCGGGTGTACAGAGTCTCGCGCGCGCTCGCGAGTACTCGAAGATCCCT 719
DB 661 CTTGACCCAGCGGGTGTACAGAGTCTCGCGCGCGCTCGCGAGTACTCGAAGATCCCT 720
QY 720 GACGCGCTCGTCAAGTTCTCCGTCGACGACAGATAAGAGAGATCTATAACGAGGATGC 779
DB 721 GACGCGCTCGTCAAGTTCTCCGTCGACGACAGATAAGAGAGATCTATAACGAGGATGC 780
QY 780 CCGGTCGTCAGTGTCCGTCCTCGACGCGCCGACGTGAGGCTCTCGATTTTC 839
DB 781 CCGGTCGTCAGTGTCCGTCCTCGACGCGCCGACGTGAGGCTCTCGATTTTC 840
QY 840 TCCTTTCCGATGTTTCAGAACCGCTTACGTCCGCGTAGCGAGCTCGAGAACAGAGATG 899
DB 841 TCCTTTCCGATGTTTCAGAACCGCTTACGTCCGCGTAGCGAGCTCGAGAACAGAGATG 900
QY 900 TCGCTCGACGTTGTCAACGACCTCATCGAGTGGCTCAACAATCTCGCGACTGGGCTTAT 959
DB 901 TCGCTCGACGTTGTCAACGACCTCATCGAGTGGCTCAACAATCTCGCGACTGGGCTTAT 960
QY 960 GTGCTGACTCTGAACAGTGGATTAACTTCAACCAATGACACACGATCTAGTCCGCAATC 1019
DB 961 GTGCTGACTCTGAACAGTGGATTAACTTCAACCAATGACACACGATCTAGTCCGCAATC 1020
QY 1020 CCGGTTCTACGTCGAACCTTACGCTTCCAGACCCCAACAGAGGCTTGTTCGACAGTC 1079
DB 1021 CCGGTTCTACGTCGAACCTTACGCTTCCAGACCCCAACAGAGGCTTGTTCGACAGTC 1080
QY 1080 TCAGACTACCGCTTCACTTATAAGCGGATAACATGTGAAGCAACATGCAACATCTGTC 1139
DB 1081 TCAGACTACCGCTTCACTTATAAGCGGATAACATGTGAAGCAACATGCAACATCTGTC 1140
QY 1140 GACCAAGGCTTTGGATCGGCGCCAGTACGCTCTCACCCGACTAGCTTACCGCAGTAC 1199
DB 1141 GACCAAGGCTTTGGATCGGCGCCAGTACGCTCTCACCCGACTAGCTTACCGCAGTAC 1200
QY 1200 GAGTACGAGGCTTACGCTTGTGACACTTTGACCTTTCGCGACGACATCGAGCGCGCT 1259
DB 1201 GAGTACGAGGCTTACGCTTGTGACACTTTGACCTTTCGCGACGACATCGAGCGCGCT 1260
QY 1260 GCATCTCGGTTGTGTGGGAGTTTGCACAGGGTGGCACTGCGCTTCGAGGCACTCCA 1319
DB 1261 GCATCTCGGTTGTGTGGGAGTTTGCACAGGGTGGCACTGCGCTTCGAGGCACTCCA 1320
QY 1320 GCCTGGGAGCAGGCACTCTCGGTTGGCTTACCTCACCTGGCGCCCAACAGGTACTATTTC 1379
DB 1321 GCCTGGGAGCAGGCACTCTCGGTTGGCTTACCTCACCTGGCGCCCAACAGGTACTATTTC 1380
QY 1380 CAGTGTGCTCGTTAGTCTGCTTCCCTGAGGGTTTCGCGCTTTCGAGGCTTACGACCG 1439
DB 1381 CAGTGTGCTCGTTAGTCTGCTTCCCTGAGGGTTTCGCGCTTTCGAGGCTTACGACCG 1440
QY 1440 AACGACGCTCTTGACCGGCTTCCGTCGAGGAGACCGTCACTTTTCGCGAGGTC 1499
DB 1441 AACGACGCTCTTGACCGGCTTCCGTCGAGGAGACCGTCACTTTTCGCGAGGTC 1500
QY 1500 GCGGTCGACGAGGTCTGTTGTGACCAACAAACCCCGCGCGGCGAGCGCCCGACCTTTC 1559

Db 1501 GCGGTGAGAGTGGTGTGACCAACACCCGCGCGGCGGAGGCGCCCACTTC 1560
QY 1560 ACCGTGAGAGTGGCCCTTAAACGGTTACACACACCGTGTGTAGGAACAGCTCTTA 1619
Db 1561 ACCGTGAGAGTGGCCCTTAAACGGTTACACACACCGTGTGTAGGAACAGCTCTTA 1620
QY 1620 GAGACTGAGACCTCTCTCGTAGGCTGGAACCTCCCTATGCCACCTGCTGCTGACAG 1679
Db 1621 GAGACTGAGACCTCTCTCGTAGGCTGGAACCTCCCTATGCCACCTGCTGCTGACAG 1680
QY 1680 ACCGTGCGCAACAAACCCGAGAGTACGAGAGTGGCTTTTAAAGAAACACTTGGCTGCTAT 1739
Db 1681 AGGTGCGCAACAAACCCGAGAGTACGAGAGTGGCTTTTAAAGAAACACTTGGCTGCTAT 1740
QY 1740 TTTGGTCACTCAAAATGCAACCCCGTTTCCAGCTCACGCGAGCGAGCTCTTTGGC 1799
Db 1741 TTTGGTCACTCAAAATGCAACCCCGTTTCCAGCTCACGCGAGCGAGCTCTTTGGC 1800
QY 1800 GCGCTTTCCTTCAACAAATCGGGTTATGAGCGCACACGCGACCTCCCGACTACACTGGC 1859
Db 1801 GCGCTTTCCTTCAACAAATCGGGTTATGAGCGCACACGCGACCTCCCGACTACACTGGC 1860
QY 1860 ATCCGTGACTCATTTGACACAGAACATGTCCACCGCTGTGGCCACTTCCGCTCACTCTCC 1919
Db 1861 ATCCGTGACTCATTTGACACAGAACATGTCCACCGCTGTGGCCACTTCCGCTCACTCTCC 1920
QY 1920 CACTCTGCACTATCGTCACTAAGACCTACAGGGTTGGGAAGGGGTACAGAACTGCAAC 1979
Db 1921 CACTCTGCACTATCGTCACTAAGACCTACAGGGTTGGGAAGGGGTACAGAACTGCAAC 1980
QY 1980 AGCCTTTTCGCGCAATTTGCGCACGCGGGCTCTCAAGAAATGAGAGATCTCTGCCTC 2039
Db 1981 AGCCTTTTCGCGCAATTTGCGCACGCGGGCTCTCAAGAAATGAGAGATCTCTGCCTC 2040
QY 2040 GCGGAGACTGCGCCACCGCTCTCACAGGTGTCTACCCGCGCACTGACAACTTCGCGGC 2099
Db 2041 GCGGAGACTGCGCCACCGCTCTCACAGGTGTCTACCCGCGCACTGACAACTTCGCGGC 2100
QY 2100 GCGCTTTCGCTTTCGCGCAACATGTCTCTCGTCTGAGTGGAGGCAAGTCC 2159
Db 2101 GCGCTTTCGCTTTCGCGCAACATGTCTCTCGTCTGAGTGGAGGCAAGTCC 2160
QY 2160 TCCATCATCAAGTCCGTTGGCGAGACTGCGGTGCGCGCGCTCAGTCCGCGCTCGCGAG 2219
Db 2161 TCCATCATCAAGTCCGTTGGCGAGACTGCGGTGCGCGCGCTCAGTCCGCGCTCGCGAG 2220
QY 2220 CTACCCGAGTGTAAATGAGTGTACAGGAGATTTGCGCGCTGTCCGCGCGCGCGA 2279
Db 2221 CTACCCGAGTGTAAATGAGTGTACAGGAGATTTGCGCGCTGTCCGCGCGCGCGA 2280
QY 2280 GCGCGCGCGCGCGCTCGTGCCTAATGTTGCTCGCTCGCTGTTTCGCGCTTCGTA 2339
Db 2281 GCGCGCGCGCGCGCTCGTGCCTAATGTTGCTCGCTCGCTGTTTCGCGCTTCGTA 2340
QY 2340 AACGGGTGTCGCGCACATTAGCGGTACCTTAAGACTCTGCTGAGTCCCGTCTGTTAC 2399
Db 2341 AACGGGTGTCGCGCACATTAGCGGTACCTTAAGACTCTGCTGAGTCCCGTCTGTTAC 2400
QY 2400 ACAGCGGTGTCGCGCGGTTCGATTCCATTTCCAGCGGCAAGAGAGTGTAGTCTC 2459
Db 2401 ACAGCGGTGTCGCGCGGTTCGATTCCATTTCCAGCGGCAAGAGAGTGTAGTCTC 2460
QY 2460 TCGTCCCTCGGGATACCA 2478
Db 2461 TCGTCCCTCGGGATACCA 2479

RESULT 4

US-09-991-262-39

; Sequence 39, Application US/09991262

; Publication No. US20030041349A1

; GENERAL INFORMATION:

; APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and Their Uses in
; Protecting Plants

; NUMBER OF SEQUENCES: 57

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fiehr Hobbach Test Albritton & Herbert LLP

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: United States

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/991,262

; FILING DATE: 20-NO. US20030041349A1-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/234,238

; FILING DATE: 20-JAN-1999

; APPLICATION NUMBER: US 08/485,355

; FILING DATE: 07-JUN-1995

; APPLICATION NUMBER: US 08/440,522

; FILING DATE: 12-MAY-1995

; APPLICATION NUMBER: US 08/089,372

; FILING DATE: 08-JUL-1993

; APPLICATION NUMBER: AU PL4081/92

; FILING DATE: 14-AUG-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Trecartin, Richard F.

; REGISTRATION NUMBER: 31,801

; REFERENCE/DOCKET NUMBER: A-58631-3/RFT/DSS

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; TELEX: 910 277299

; INFORMATION FOR SEQ ID NO: 39:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5312 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 37..5148

; SEQUENCE DESCRIPTION: SEQ ID NO: 39:

US-09-991-262-39

Query Match 2.9%; Score 71.2; DB 11; Length 5312;

Best Local Similarity 78.7%; Pred. No. 2e-10;

Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 2371 TAAAGACTCTGGTGTGATCCCGTCTACAGACGGGTCTGCGGTTTCGATTCCATTTC 2430

Db 5205 TAAAGACTCTGGTGTGATCCCGTCTACAGACGGGTCTGCGGTTTCGATTCCATTTC 5264

QY 2431 CCAAGCGGCAAGAGAGTGTAGTCTGCGTCCCTCGGGATACCA 2478

Db 5265 CCAAGCGGCAAGAGTGTAGTCTGCGTCCCTCGGGATACCA 5312

RESULT 5

US-09-991-262-41

; Sequence 41, Application US/09991262

; Publication No. US20030041349A1

; GENERAL INFORMATION:

; APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.

; TITLE OF INVENTION: Insect Viruses and Their Uses in

; Protecting Plants


```

; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/991,262
; FILING DATE: 20-NOV-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,238
; FILING DATE: 20-JAN-1999
; APPLICATION NUMBER: US 08/485,355
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/440,522
; FILING DATE: 12-MAY-1995
; APPLICATION NUMBER: US 08/089,372
; FILING DATE: 08-JUL-1993
; APPLICATION NUMBER: AU PL4081/92
; FILING DATE: 14-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard P.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-586311-3/RFT/DSS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELE: 910 277299
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5368 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4944..5162
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-991-262-45

```

```

Query Match      2.9%; Score 71.2; DB 11; Length 5368;
Best Local Similarity 78.7%; Pred. No. 2e-10;
Matches 85; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 2371 TAAAGACTCTGGTGGTCCCGCTTACACGACGGGTCTGCGCGGTTTCGATTCCATTTC 2430
DB 5261 TAAAGACTCTGGTGGTCCCGCTTACACGACGGGTCTGCGCGGTTTCGATTCCATTTC 5320

QY 2431 CCAAGCGCAAGAGGACGTAGTCTTGGCTCCCTCGGGATACCA 2478
DB 5321 CCAAGCGCAAGAGGACGTAGTCTTGGCTCCCTCGGGATACCA 5368

```

```

RESULT 8
US-10-156-761-6810
; Sequence 6810, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089

```

```

; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 6810
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(750)
; US-10-156-761-6810

Query Match      2.0%; Score 48.4; DB 15; Length 750;
Best Local Similarity 46.2%; Pred. No. 0.00069;
Matches 197; Conservative 0; Mismatches 226; Indels 3; Gaps 1;

QY 1796 TGGCGCCGTTTCCTTCAACAATCCGGGTTATGAGCGCACACGCGACCTCCCGGATACAC 1855
DB 285 TGTGCGGTGCTCATCGAGTGTCTCACCGAACACCGCGCGGCTTCGGACGTCG 344

QY 1856 TGGCATCCGTGACTCATTCGACCAAGATGTCACCGCTGTGGCCAC---TTCGGTTC 1912
DB 345 CGTCGCCATGACCCCGCAACGCGCGGCAACATGGCCGACCGGGCTCCGTTCGTACTCTT 404

QY 1913 ACTCTCCCACTCTCGAGTATCGTCACTAAGACCTTACCAGGGTTGGGAAGCGGTCAACGAA 1972
DB 405 CAACCGCAAGGGCGTGTGATCGTCCCAAGGGCGAGCTGAGCGAGGACGACGTCTCTGGG 464

QY 1973 CGTCAACACGCTTTTCGCCCAATTCCGCGCACGCGGGCTCTCTCAAGAATGAGGATCTT 2032
DB 465 TGCCTGTCTCGACGCGGGCGCGGAGGAGTCAACGACCTCGGTGAGTCTCTCGAGGTCT 524

QY 2033 CTGCTCTCGCGACGACCTGGCCACCGGTCTCACAGGTCTTACCCCGCCACTGACAATT 2092
DB 525 CTCGAGGCGCACCGACCTGGTTCGCGGTCCGACCGCCCTCCAGGACGCGCGGATCGACTA 584

QY 2093 CGCGCGCCCGTTTCTGCTTCCGCGCAACATGTCTCTCGGTCTCGTGAAGTCCGAGGC 2152
DB 585 CGACTCGCGGACGCGCAACTTCTGTCGCGACCATGAGCTCGAGCTCGACGAGGCGCG 644

QY 2153 AACGTCCTCCATCATCAAGTCCGTTGGCGAGACTCCGCTCGCGCGGCTCAGTCGGCT 2212
DB 645 CGGGAAGATCTTCAAGCTGATCGACGCGCTCGAGGACAGCGACGCTCGAGACGTCCT 704

QY 2213 CGCGAA 2218
DB 705 CGCGAA 710

RESULT 9
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608

```

```
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match      2.0%; Score 48.4; DB 15; Length 9025608;
Best Local Similarity 46.2%; Pred. No. 0.02;
Matches 197; Conservative 0; Mismatches 226; Indels 3; Gaps 1;

QY 1796 TGGCGCGTTCCTCAACAATCCGGGTATGAGCGCACACCGGACTCCCGGACTAC 1855
Db 8155983 TGTCCGGTGTCTCATCGAGTGTCTCACCGACAACCGCAACCGCGCTCGGAGCTCG 8156042
QY 1856 TGGCATCCGTGACTCATTCGACAGAAATGTCACCGGTGTGGGCCAC-...TTCGGTTC 1912
Db 8156043 CGTCCCATGACCCGCAACCGGGGCAACATGGCCGACCGCGGCTCGTCTGTACTCTTT 8156102
QY 1913 ACTCTCCCACTCTCTGAGTATCGTCACTAAGACCTTACCAGGGTGTGGGAAGCGTCAAGAA 1972
Db 8156103 CAACCGCAAGGCGTGTGATCGTCCCAAGGGCGAGCTGAGCGAGGACGAGTCTCTGG 8156162
QY 1973 CGTCAACAGCCCTTTCCGGCAATTCGCGCACCGGGGCTCTCAAGATGAGGATCGT 2032
Db 8156163 TGCCGTGTCTGACGCGCGGCGCGAGGAGTCAACGACCTCGGTGAGTCTTTCGAGGTCT 8156222
QY 2033 CTGCTCGCGGACGACCTTGCGCACCGCTCTACAGGTGTCTACCGCGGCTGACAACTT 2092
Db 8156223 CTCCGAGGCCACCGACCTGTTCGGGTTCGCGGTCGCGACCGGCTTCCAGGACCGCGGATCGACTA 8156282
QY 2093 CGCGCGCGCGTTCCTTGCCTTCGCGCGCAACATGTGTCTCTCGTCTGAAGTCGGAGGC 2152
Db 8156283 CGACTCGCGGAGCGCAACTTCTGTCGCGACCATGAGTCTGAGTCTGAGCGAGGCGC 8156342
QY 2153 AAGTCTCCATCATCAAGTTCGTTGGGAGACTGCGGTGCGCGGCGGTCTAGTTCGGGCT 2212
Db 8156343 CGGGAAGATCTTCAAGCTGATCGACGCGCTCGAGGACGACGACGAGTGTCTT 8156402
QY 2213 CGCGAA 2218
Db 8156403 CGCCAA 8156408

RESULT 10
US-10-156-761-6911
; Sequence 6911, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: HORIKAWA, JUN
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 6911
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1416)
US-10-156-761-2231
Query Match      1.9%; Score 47.8; DB 15; Length 1926;
Best Local Similarity 48.4%; Pred. No. 0.0015;
Matches 133; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 1877 CCAGAACATGTCACCGCTGTGGGCCACTTCCGCTCTCTCCCACTCTCTGAGTATCGT 1936
Db 315 CGAGCACGACGTATCGAGAACTCGACGCTCTCAACGCTCTCGGCTGGGCGGCGCT 374
QY 1937 CACTAAGACCTTACGAGGTTCGGAGGCGTCAAGACCTCAACGCGCTTTTCGGCAATT 1996
Db 375 CGCCAAGGCCAACCGAGGTGCTCGAAGCGGACGACCGTGTGCGGTCTATCGGCGCG 434

US-10-156-761-6911
Query Match      1.9%; Score 48.2; DB 15; Length 1416;
Best Local Similarity 45.9%; Pred. No. 0.001;
Matches 164; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

QY 1825 ATGAGCGCACACGCGACCTCCGGACTACACTGGCATCCGTGACTCATTCGACCAAGAACA 1884
Db 215 ACAGATCGAGCGGCACCTCTTCGACGCGCGAGGAGTTCGACACGCGAGTCCCGCA 274
QY 1885 TGTCCACCGCTGTGGCCCACTTCGCTCACTCTCCACTCTCTGCAAGTATCGTCACTAAGA 1944
Db 275 CGGGCTCGCCCTCGCGTCTCTGGGACCTTCGCTCTAGCGGAGGCTCGACGGCGCTCG 334
QY 1945 CTTACACGGGTGGGAAGCGTCAAGAACTGTCGCAAGCTCAACACAGCCCTTTCGGCCAAATTCGCGCAGC 2004
Db 335 CGTGTCTCGCAGGTACGCGCGCTGCGGCAACCACTGGGCGCTTGGGCGCTCGACGAGTGG 394
QY 2005 CGGGCTCTCTCAAGATGAGGAGATCTCTGCTGCTGCGCGGCGGCTTTCGCGCAGCAGCTGCTCA 2064
Db 395 CGCTCGCGGACGACGACGCGGCGCTTGGCGGCGCTTGGCGGCTTGGCGGCGCTTTC 454
QY 2065 CAGGTGTCTACCCGCGCACTGACAACTTTCGCGGCGCGGCTTTCGCTTTCGCGCGAACA 2124
Db 455 CGTGTGATGCGAGGCGGCGGAGCTGCGCGCGCGCTTGGCGGCGGCTTTCGAAACCC 514
QY 2125 TGCTGTCTCTCGTGTGAAGTCGAGGCAACGCTCTCTCATCATCAAGTTCGTTGCG 2181
Db 515 GGCCTTGGCGGCTGTGGGCGGAGGATCGCGGAATCCATCGGACGCGGCTGGCG 571

RESULT 11
US-10-156-761-2231
; Sequence 2231, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 2231
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
US-10-156-761-2231
Query Match      1.9%; Score 47.8; DB 15; Length 1926;
Best Local Similarity 48.4%; Pred. No. 0.0015;
Matches 133; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 1877 CCAGAACATGTCACCGCTGTGGGCCACTTCCGCTCTCTCCCACTCTCTGAGTATCGT 1936
Db 315 CGAGCACGACGTATCGAGAACTCGACGCTCTCAACGCTCTCGGCTGGGCGGCGCT 374
QY 1937 CACTAAGACCTTACGAGGTTCGGAGGCGTCAAGACCTCAACGCGCTTTTCGGCAATT 1996
Db 375 CGCCAAGGCCAACCGAGGTGCTCGAAGCGGACGACCGTGTGCGGTCTATCGGCGCG 434
```

QY 1997 CGCGACGCGGGCTCTCAAGATGAGGAGATCTCTGCTCGCGACGACCTGGCCAC 2056
Db 435 CGCCCTCACGGCGGATCGCTGGGAGCGCTGAACACATCGCGGCGCAAGACG 494
QY 2057 CGGTCTCACAGGTGTCTACCGGCGACTGACAACTTCGGCGGCGCGGTTCTGCTTGC 2116
Db 495 CGCCCTCGTCTGTCGTGTAACGACAAACAGGCGCTGTACGCGCGCCACGATCGGCGGCT 554
QY 2117 CGCGACATGCTGCTCTCGTCTGTAAGTCGGAGG 2151
Db 555 CGCCAACCACTTGGGCCACCTTGGCACCGGACG 589

RESULT 12

US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 1.9%; Score 47.8; DB 15; Length 9025608;

Best Local Similarity 48.4%; Pred. No. 0.03;
Matches 133; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 1877 CCAGAACATGTCCACCGCTGTGGCCCACTTCCGCTCACTCTCCCACTCTCTGCAAGTATCGT 1936
Db 2732084 CGAGCAGCGCTATCGAGAACTCGACGCTTCCAGGCTCTCGGCTGGGCCGACGGCCT 2732025
QY 1937 CACTAAGACCTACAGGGTTGGAAGCGGTACGAAAGTCAAGATGAGGATCTCTGCTCGCGACGACCTGGCCAC 2056
Db 2732024 CGCAAGGCCAACACAGGTGTCTGAAACGCGACGACGCTGGTGGCGGTATCATCGGCGACG 2731965
QY 1997 CGCGACGCGGGCTCTCAAGATGAGGAGATCTCTGCTCGCGACGACCTGGCCAC 2056
Db 2731964 CGCCCTCACGGGCGATGCGCTTGGGAGCGCTGAAACAACTCGCGCGCGCCAAAGACCG 2731905
QY 2057 CGGTCTCACAGGTGTCTACCCCGCACTGACAACTTCGGCGGCGCGGTTTCTGCTTGC 2116
Db 2731904 CCCCCTCGTATCGTGTGAAACACAGCGCTGTAGCGCCCGCCACGATCGGCGGCT 2731845
QY 2117 CGCGAACATGCTGTCTCGTGTGTAAGTCGGAGG 2151
Db 2731844 CGCAACCACTTGGCCACCTTGGCACCGGACG 2731810

RESULT 13

US-10-156-761-6645
; Sequence 6645, Application US/10156761
; Publication No. US20030119018A1

; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 6645
; LENGTH: 3570
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3570)
US-10-156-761-6645
Query Match 1.9%; Score 47; DB 15; Length 3570;
Best Local Similarity 45.9%; Pred. No. 0.0032;
Matches 161; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
QY 1880 GAACATGTCCACCGCTGTGGCCCACTTCCGCTCACTCTCCCACTCTCTGCAAGTATCGTCA 1939
Db 1107 GTACATGCGATCGGTGAGCGGCACCAACGCGCTCCAAAGATTCCGCGAGGCCAT 1166
QY 1940 TAAGACCTACAGGGTTGGGAGGCGTCAAGAACCTCAACACGCTTTTCGGCCAAATTGCG 1999
Db 1167 GCTGACCGCCCGCTGGACGACTCGTGGAGATGCGCGCGACCAAGATCCGCGAGGGCGC 1226
QY 2000 GCACGCGGGCTCTCTCAAGATGAGGAGATCTCTGCTCTCGCGCGACGACTGGCCACCGC 2059
Db 1227 GCACATGCTGACCTGTGGTGCATACGTGGCGCGGCGGCTCGCGACATGGAGGA 1286
QY 2060 TCTCAGGTGTCTACCCCGCCACTGACAACTTTCGGCGCGCGGCTTCTGCTTGGCCCGC 2119
Db 1287 ACTGCGCGCGCTTTCGCCACCGCTCCACGCTGCGGATCGTGTGGAATCTCCACCGAGGT 1346
QY 2120 GAACATGCTGCTCTCGTGTGTAAGTCGAGGCAAGCTCTCCATCATCAAGTCCGTTGG 2179
Db 1347 CGACGCTCTGCGCGCGGACTGGAGAGCTCGCGCGCGCGGCTCATCAACTCCGTCAA 1406
QY 2180 CGAGACTGCGCGCGGCTCAGTTCGGGCTTCGCGAGCTTACCGGACT 2230
Db 1407 CTACGAGGACGGGACGGCCCGAGTTCGGGTTTCGGGAAGGTCAACCGGCT 1457

RESULT 14

US-10-156-761-176
; Sequence 176, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30

```

: PRIOR APPLICATION NUMBER: JP 2001-272697
:
: PRIOR FILING DATE: 2001-08-02
:
: NUMBER OF SEQ ID NOS: 15109
:
: SEQ ID NO 176
:   LENGTH: 2679
:   TYPE: DNA
:
: ORGANISM: Streptomyces avermitilis
:
: FEATURE:
:   NAME/KEY: CDS
:   LOCATION: (1)..(2679)
US-10-156-761-176

```

Query Match	1.9%;	Score 46.4;	DB 15;	Length 2679;
Best Local Similarity	50.0%;	Pred. No. 0.0044;		
Matches 116;	Conservative 0;	Mismatches 116;	Indels 0;	Gaps 0;
Qy	2002	ACGCGGCGCTCCTCAAGAATGAGAGATCCTTCGCTCGCGAGCAGCTGGCCACACCGTC	2061	
Db	1028	ACCAGGTCCGGCGCAAGAAGACGAATGGTCGACAGGCCGACCACTGACGCGCCGCC	1087	
Qy	2062	TCACAGTGTCTACCCCGCCACTGACAACTTCGGCGCGCGCGTTTCCTTCGCGCGGCA	2121	
Db	1088	CCACCGGATGCGCGCTCCTCGTACCGATGCGCGCCGCGACCTGGCGAGCTGGACGCC	1147	
Qy	2122	ACATGCTGTCTCGTGCTGAAGTCGAGGCAAGCTCTCATCATCAAGTCGGTTGCGG	2181	
Db	1148	GCCTGTCGCGTACGTTCGTCGAGTCCAGTCCGTACCGTCTGACCGGTGTTGGG	1207	
Qy	2182	AGACTCCGTCGGCGCGGCTCAGTCCGGCTTCGCAAGTACCCGGACTGCT	2233	
Db	1208	GGCGCACCGCGCGCGTGTGGACGCCCGCGCCCGGCGCTTCGGCGCGCTGCT	1259	

```

RESULT 15
US-10-156-761-5243
; Sequence 5243, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 5243
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(480)
US-10-156-761-5243

```

	Query Match	1.9%	Score 46.2;	DB 15;	Length 480;
	Best Local Similarity	52.9%	Pred. No. 0.0027;		
	Matches 99;	Conservative 0;	Mismatches 88;	Indels 0;	Gaps 0;
Qy	1413	GGTTTCGCCCTTGAGCGCTACGACCCGAAAGAGCGGCTCTTGGACCGACTTCGCTTCGCA	1472		
Db	202	GGGTACGCCCTGTGGAAACGGCCGACCCCGAGCGGCGCCACCCCGGCGGACTCCGTCCGGGCCC	261		
Qy	1473	GGAGACACCGTCACTTTTCGGCGAGGTGCGCGTGGACGAGGTGCTGTGTGACCAACAACCCC	1532		

Db	262	TGCACACGACGACCGACGGGTGGGGCATCGAGGCCCGGCTCGACATCAACCGGACCGC	321
Qy	1533	GC CGCGCGCGCAGCGGCCCCACCTTTCACCGTGAGAGTGCCCCCTTCAAAAGCTTACACC	1592
Db	322	ACCATCGACGCGACCGGCTCCACCGCGGGCGGACCGCGCGTACTGCACCGCGTGAAG	381
Qy	1593	AACACCG	1599
Db	382	AGCGGCG	388

Search completed: January 13, 2004, 01:21:55
Job time : 865 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 18:17:41 ; Search time 5103 Seconds

(without alignments)
11802.174 Million cell updates/sec

Title: US-09-991-262-47

Perfect score: 2478

Sequence: 1 GTTTTCTTTCTTTACCAAG.....CTGGCTCCTCGGATACCA 2478

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_ges_hum:*

18: em_ges_inv:*

19: em_ges_pln:*

20: em_ges_vrt:*

21: em_ges_fun:*

22: em_ges_mam:*

23: em_ges_mus:*

24: em_ges_pro:*

25: em_ges_rod:*

26: em_ges_phg:*

27: em_ges_vrl:*

28: gb_gsel:*

29: gb_gsel2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	58.6	2.4	766	29	AG090630 Pan trogl
2	54.2	2.2	638	10	B8806756 ss65a07.y
c 3	53.4	2.2	693	14	CB685967 OSJNEF16N
4	52.4	2.1	925	29	CNS0091P

5	51	2.1	578	13	BQ742011
6	51	2.1	637	14	CA754508
c 7	50.4	2.0	1540	29	AG149328
8	50.2	2.0	689	10	BE440409
9	50.2	2.0	689	14	CB685120
10	50.2	2.0	691	10	BF430707
11	50.2	2.0	705	10	BF430789
c 12	50.2	2.0	710	14	CB685121
13	50.2	2.0	776	14	CB685782
c 14	50.2	2.0	781	14	CB651317
c 15	50.2	2.0	793	14	CB666855
16	50.2	2.0	796	14	CB685966
17	50.2	2.0	804	12	BI305514
18	50.2	2.0	816	14	CB684015
c 19	50.2	2.0	821	14	CB684016
20	50.2	2.0	829	14	CB666854
21	50.2	2.0	830	14	CB666607
c 22	50.2	2.0	836	14	CB666608
23	50.2	2.0	853	14	CB681881
24	50.2	2.0	854	14	CB685905
25	50.2	2.0	882	14	CB651316
26	50	2.0	437	12	BM134400
27	49.6	2.0	584	12	BM525740
c 28	49	2.0	925	29	CNS0091P
29	48.6	2.0	519	14	CB212054
30	48.6	2.0	627	12	BM137775
c 31	48.6	2.0	651	14	CB618894
c 32	48.6	2.0	664	14	CB632421
33	48.6	2.0	700	10	BF430538
c 34	48.6	2.0	709	14	CB632274
35	48.6	2.0	754	14	CB618893
36	48.6	2.0	826	14	CB632420
37	48.6	2.0	849	14	CB632273
c 38	48.6	2.0	1000	13	BM407619
39	48.4	2.0	571	14	CB685925
40	48	1.9	562	9	AM620815
41	48	1.9	765	13	BQ802457
c 42	47.8	1.9	1201	13	BM376097
43	47.4	1.9	902	29	CC349811
44	46.6	1.9	538	14	CA599587
c 45	46.6	1.9	885	13	BM425603

ALIGNMENTS

RESULT 1	AG090630/c	AG090630	766 bp	DNA	linear	GSS 03-NOV-2001
LOCUS	Pan troglodytes DNA, clone: PTB-090D23.R, genomic survey sequence.					
DEFINITION	Pan troglodytes DNA, clone: PTB-090D23.R, genomic survey sequence.					
ACCESSION	AG090630					
VERSION	AG090630.1	GI:16642432				
KEYWORDS	GSS.					
SOURCE	Pan troglodytes (chimpanzee)					
ORGANISM	Pan troglodytes					
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.					
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.					
TITLE	BAC end sequences of Library PTB					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 766)					
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.					
TITLE	Direct Submission					
JOURNAL	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Sushiro-chou, Tsukuba, Ibaraki, Japan 305-8565, Japan (E-mail: chimps@gsr.riken.go.jp, URL: http://hgp.gsfc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)					
COMMENT	Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of					

```
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..766
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/cloned="PTB-090D23.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
23 a 42 c 352 g 11 t 338 others
BASE COUNT
ORIGIN
Query Match 2.4%; Score 58.6; DB 29; Length 766;
Best Local Similarity 25.7%; Pred. No. 0.00012;
Matches 184; Conservative 0; Mismatches 533; Indels 0; Gaps 0;
QY 1499 CGCGGTGACGAGGTGTTGTGACCAACACCCCGCGCGCGGCGAGCGCCCGCCACCTT 1558
Db 719 CCCCCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
QY 1559 CACCGTGAGAGTGCCCGCTTAAACGCTTACACCAACACGCTGTTTGTAGGAAACAGCTCTT 1618
Db 659 NNNCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 600
QY 1619 AGAGACTCGACCTCTCTCGTAGGCTCGAATCCCTATGCCACCTGCTGACTTTGGACA 1678
Db 599 CNCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 540
QY 1679 GACGCTGCGCAACCAACCCGGAAGATCGAGCTGCTTCTTAAAGAAACACTTGGTGCTA 1738
Db 539 CNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
QY 1739 TTTGGTGCATCTCAAAATGCGAAACCCCGTTTTCCAGCTACGCGAGCAGCTCTTTGG 1798
Db 479 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420
QY 1799 CGCGGTTCTTCAACAATCCGGGTTATGAGCGCACACGCGACTCCCGGACTACACTG 1858
Db 419 CNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 360
QY 1859 CATCCGTGACTATTGCAACGAGAAACATGTCCACCGCTGTGGCCCACTTCGGCTCACTTC 1918
Db 359 CCCCCCCCCCCCCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
QY 1919 CCACTCCTGAGTATCGTCACTAAGACTTACAGGTTGGGAAAGGCGTCAAGCAAGTCAA 1978
Db 299 CCCCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
QY 1979 CAGCCTTCGGCAATTGCGGACGCGGCTCTCTCAAGATGAGAGATCCTCTGCT 2038
Db 239 CCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180
QY 2039 CGCGGAGAGCTGGCCACCGCTCTCAGAGTGTCTACCGCGGCACTGCAACTTCGCGGC 2098
Db 179 CNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 120
QY 2099 CGCGGTTTCGCTTCGCGCGCAACATGTGTCTCTCGTGTGCTGAAGTCGAGGCAAGTC 2158
Db 119 CNCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 60
QY 2159 CTCATCATCACTCGTTGGCGGAGCTGCGGCGGCTCAGTTCGCGGCTCGC 2215
Db 59 GNCNNNNNNAGTTCTCTTTTATATGCGGGGGGGTGGCGCCNCCNCCNCCNCCNCCNCC 3
```

```
BE806756 638 bp mRNA linear EST 06-DEC-2001
LOCUS ss65a07.v1 Gm-cl062 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl062-1261 5' similar to TR:P93726 P93726 PEROXIDASE ATP26A ;
mRNA sequence.
ACCESSION BE806756
VERSION BE806756.1 GI:10237868
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 638)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurr
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 410.
Location/Qualifiers
1..638
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl062-1261"
/tissue_type="stem tissue of greenhouse grown plants"
/dev_stages="1 month old"
/lab_host="DH10B"
/clone_lib="Gm-cl062"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from stem tissue of 1 month old greenhouse grown plants
for the cultivar Raiden. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy Shoemaker."
BASE COUNT 116 a 292 c 109 g 120 t
ORIGIN
Query Match 2.2%; Score 54.2; DB 10; Length 638;
Best Local Similarity 50.6%; Pred. No. 0.0019;
Matches 131; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
QY 1378 TCCGAGTCGTCCTGTTAGCTTCCCTTCCCTGAGGTTTCCCTTGGAGCTGACGACC 1437
Db 178 TCCCGGCGACCCCTTCGACCTCGTCGTCGCGCAAGACCCGCTCTCGAGCTCTCTGCG 237
QY 1438 CGAACGACCGCTCTTGGACCGGACTTCGCTTCGCGAGGAGACACCGTCACTTTCCGCGAG 1497
Db 238 CGAACACCGTCTCTGCTCCGACATCTCTCCGCGCGCACCCGCGACCTCTCTCACCATGC 297
QY 1498 TCGCGCTCGACGAGGTGCTGTGTGACCAACAACCCCGCGCGCGGCGGCGGCCCCACCT 1557
```

```
FEATURES
source
```

```

Db      298  TCGGGGGCCCTTCTTCCCGCTTCTTCTCGGGCGCGGACGGCGAACCTCCCTCGCCT 357
QY      1558  TCACCGTGAGAGTGGCCCCCTTCAACAGCTTACACCAACACCGTGTGTTAGGAACAGCTCT 1617
Db      358  CGCGGTATCTCTCCACCTCCCGCCCTCCATGCCATCTCTCAATCACCACCTCT 417
QY      1618  TAGAGACTCGACCTCTCTC 1636
Db      418  TCGCCAACGCGGCTTCAC 436

RESULT 3
LOCUS   CB685967/c
DEFINITION OSJNEf16N13.r OSJNEf Oryza sativa (japonica cultivar-group) CDNA
ACCESSION CB685967
VERSION   CB685967
KEYWORDS  CB685967.1 GI:29689692
SOURCE   EST.
ORGANISM Oryza sativa (japonica cultivar-group)
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
          Ehrhartoideae; Oryzae; Oryza.
REFERENCE 1 (bases 1 to 693)
AUTHORS  Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
          Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE    Large-scale identification of ESTs involved in the interaction
JOURNAL  Unpublished
COMMENT  Contact: Rod Wing
          Arizona Genomics Institute
          University of Arizona
          Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
          85721-0088, USA
          Tel: 520 626 3967
          Fax: 520 621 9288
          Email: http://genome.arizona.edu
          PCR Primers
          FORWARD: gta aaa cga cgg cca gtg
          BACKWARD: gga aac agc tat gac cat g
          Plate: 16 row: N column: 13
          Seq primer: gga aac agc tat gac cat g.

FEATURES             source
1..693
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/db_xref="Nipponbare"
/clone="OSJNEf16N13"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEf"
/notes="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control"

BASE COUNT    143 a   200 c   236 g   114 t
ORIGIN
Query Match      2.2%; Score 53.4; DB 14; Length 693;
Best Local Similarity 51.5%; Pred. No. 0.0031;
Matches 123; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY      1958  GGAAGCGTCAGACGCTCAACACGCTTTCGGCCAAATTCGGCGACGGGGCTCCTCA 2017
Db      498  GTCTGCGCTTCGCCACGCTACTCTCTGTCTGGGGCTTTCGGCTACCGGCGCGCTCG 439
QY      2018  GAATGAGGAGATCTCTGCTCGCGACGACCTGGCCACCCGCTCTCACAGGTGTCTACCC 2077
Db      438  CAAACGCGGTGACAGCGGACGCGCGCCCTTCGGCGCGGAGGCTCTGCGCCATCTCT 379
QY      2078  CCCCCTGACAACTTCGGGGCGCGCTTCTGCTCTGCGCGGAAACATGCTCTCTCCGT 2137

```

```

Db      378  CAGCGTTTGCCTCACCATGTAGCGGCTCGCCTCTTTCGGCGAGGCGGAGCGCTCCACGGC 319
QY      2138  GCTGAAGTCGGAGGCAACGCTCTCCATCATCAAGTCCGTTGGCGAGAGCTGCGTGGCGG 2196
Db      318  GCCGAGCTGAGCTCACCGGCGGCAAGAGGAGCGCCACAGCTGCAGACGCGCGCAGC 260

RESULT 4
LOCUS   CNS0091P
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
          BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
          fly), genomic survey sequence.
ACCESSION AL053013
VERSION   AL053013
KEYWORDS  CNS0091P
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
AUTHORS  Genoscope.
JOURNAL  Direct Submission
COMMENT  Submitted (02-JUN-1999) Genoscope - Centre National de Sequenage :
          BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
          - Web : www.genoscope.cns.fr)
          Determination of this BAC-end sequence was carried out as part of a
          collaboration with the Berkeley Drosophila Genome Project (BDGP).
          The BDGP is constructing a physical map of the Drosophila
          melanogaster genome using these BACs. For further information
          please see http://www.fruitfly.org The BDGP Drosophila
          melanogaster BAC library was prepared by Kazutoyo Osoegawa and
          Aaron Mammosser in Pieter de Jong's laboratory in the Department of
          Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
          NY. The library is named RPCI-98 and was constructed by partial
          EcoRI digestion of Drosophila DNA provided by the BDGP from the
          isogenic strain y2; cn bw sp, the same strain used for the BDGP's
          p1 and EST libraries. A more detailed description of the library
          and how to order individual BAC clones, the entire library, or
          filters for hybridization from the BACPAC Resource Center can be
          found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES             source
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/notes="end : TET3"

BASE COUNT    120 a   61 c   61 g   172 t   511 others
ORIGIN
Query Match      2.1%; Score 52.4; DB 29; Length 925;
Best Local Similarity 13.3%; Pred. No. 0.0061;
Matches 52; Conservative 173; Mismatches 165; Indels 0; Gaps 0;

QY      1882  ACATGTCCACCGCTGTGGCCACTTCGCTCACTCCACTCTCTGCACTATCGCACTA 1941
Db      536  MCGKCGTGTGSGTGTGTTTSSGSGYKCGCSGSGBSGCCSCSCSSSCSCBCCCCCS 595
QY      1942  AGACCTACGAGGTGGGAAGCGGTACGAACGCTCAACACGCTTTCGGCCAAATTCGGCG 2001
Db      596  CSSYCCSSBSBSSKCSSTBSBCSCCCSKSVCGTSCSSSSSCSSSSSTSSSTSSSK 655
QY      2002  ACGCGGGCTCTCAAGAAATGAGGAGATCTCTGCTCGCGGACGACCTGCGCACCGTTC 2061
Db      656  SSSSSSSSSSYTTSTKSTASGSGWSAGGSGSTGTSSTSSSSSSTSTSSSSSVSGS 715
QY      2062  TCACAGGTGTCTACCCCGCCACTGCAACTTCGCGGCGCGCTTCTGCTTCCGTCGCGCGA 2121
Db      716  KSSTBSGSGBSGSSSSSSTSSBSCTSTSSSSSSSYSSSTCSCTCTCCCTCCSYSSSTSS 775

```


Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

1 (bases 1 to 628)

REFERENCE
AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
A., Bolla,B., Maria,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project

TITLE
JOURNAL
COMMENT

Public Soybean EST Project

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Hunteville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cou@resgen.com
High quality sequence stop: 478.

FEATURES

Source

1..628
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl043-1166"
/tissue_type="Hypocotyl and plumule, germinating seeds"
/lab_host="DH10B"
/clone_lib="Gm-cl043"
/note="Vector: pT73Pac (Pharmacia); Site_1: EcoRI;
Site_2: NotI; This cDNA library was constructed from mRNA
isolated from hypocotyl and plumule tissues of seeds
germinated for three days of the cultivar Williams.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly(dT) sequence with a NotI restriction
site. EcoRI adapters were ligated to the blunt-ended cDNA
fragments followed by digestion with EcoRI and NotI. The
cDNA fragments were directionally cloned into the
EcoRI-NotI restriction site of the pT73-Pac vector. The
ligated cDNA fragments were transformed into DH10B host
cells (Gibco BRL). This library was constructed by Dr.
Randy Shoemaker."

BASE COUNT
ORIGIN

105 a 285 c 112 g 125 t 1 others

Query Match 2.0%; Score 50.2; DB 10; Length 628;

Best Local Similarity 52.4%; Pred. No. 0.023;

Matches 109; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 1378 TCCAGTGGCTCCGTTAGCTTCTCCCTGAGGGTTTCGCCCTTGAGCGGTACGACC 1437

DB 306 TCCCGGGAGCGCTTCGACTCGTCGTCGCGCCGACGCGCTTGAGTCGCTGCC 365

QY 1438 CGAACGACGCTTTGAGACGCACTTCGCTTCGCGAGAGACACGTCACCTTCGCGAGG 1497

DB 366 CCAACACCGTCTCTGCGCGACATTTCTTCGCGCGCACCGCGACCTCTCACCATGC 425

QY 1498 TCGCGCTCGACGAGGTGTTGTGACCAACACCCCGCGCGCGAGCGCCCGCCACCT 1557

DB 426 TCGCGCGCCCTTCTTCCCGCTTCTTCGCGCGCGCGAGCGCGACCTCTCGCT 485

QY 1558 TCACCGTGAGGTGCCCGCTTCAACGC 1585

DB 486 NCGCGTTCCTGACCACTCCCGCCACCC 513

RESULT 9

CB685120

LOCUS

OSJNEf15E15.f OSJNEf Oryza sativa (japonica cultivar-group). cDNA
linear EST 09-APR-2003
689 bp mRNA

clone OSJNEf15E15 5', mRNA sequence.

CB685120

ACCESSION
VERSION

CB685120.1 GI:29688845

KEYWORDS

EST.

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 689)

AUTHORS

Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

TITLE

Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea

JOURNAL

Unpublished

COMMENT

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aaa cga cgg cca gtg

BACKWARD: gga aac agc tat gac cat g

Plate: 15 row: E column: 15

Seq primer: gta aaa cga cgg cca gtg.

Location/Qualifiers

1..689

FEATURES

source

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="OSJNEf15E15"

/tissue_type="Leaf"

/dev_stage="3 week"

/lab_host="DH10B"

/clone_lib="OSJNEf"

/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control"

BASE COUNT 113 a 228 c 194 g 154 t

ORIGIN

Query Match 2.0%; Score 50.2; DB 14; Length 689;

Best Local Similarity 50.6%; Pred. No. 0.023;

Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1958 GGAAGCGCTCACGAACGTCACACACGCTTTCGGCCCAATTCGCGCAGCGGGCTCTCAA 2017

DB 126 GGTGGGCTCGCCACGGCTACTCTCTGTGGGCGCTTCGGCTCACCGGCGGCTCGG 185

QY 2018 GAATGAGGAGATCTTGTGCTTCGCGCAGACCTGGCCACCCCTCTCACAGGTGTCTACCC 2077

DB 186 CAACACGCGGTGACGCGCAGCGCGGCGCCCTTCGCGCGCGAGGCTCTGTGCCATCT 245

QY 2078 CGCCACTGACAACTTCGCGCGCGGCTTTCCTCTTCGCGCGGAACATCTGTCTCTCCGT 2137

DB 246 CAGCGTTCCTCTACCATGTACGGGTGCGCTTCTTCGCGGAGGGGAGCCGCTCCACGGC 305

QY 2138 GCTGAAGTCGAGGCAACCTCTCCATCATCAAGTCCGTTGGCGAGACTGCCGTCGGCG 2196

DB 306 GCGACGCTGACGCTCACCGCGCGGCGGAGGCGGCAAGCTGCAGACCGCGACG 364

RESULT 10

BF430707

LOCUS

OG03G0773 OG Oryza sativa cDNA clone OG03G07 5', similar to
PHOTOSYSTEM I REACTION CENTRE SUBUNIT XI PRECURSOR (SUBUNIT V),
mRNA sequence.

ACCESSION

BF430707

VERSION

BF430707.1 GI:11442808

691 bp

mRNA

linear

EST 20-FEB-2001

```

KEYWORDS      EST.
SOURCE         Oryza sativa
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
               Ehrhartoideae; Oryzeae; Oryza.
REFERENCE      1 (bases 1 to 691)
AUTHORS        Cushman, J.C. and Nomura, M.
TITLE          A cDNA library from Oryza sativa cv. Pokkali leaves enriched for
               salt-responsive genes
JOURNAL        Unpublished
COMMENT        Contact: Cushman JC
               Department of Biochemistry
               University of Nevada
               MS200, Reno, NV 89557-0014, USA
               Tel: 775-784-1918
               Fax: 775-784-1650
               Email: jcushman@unr.edu
               PCR Primers
               FORWARD: T7
               BACKWARD: T3
               Plate: OG03 row: G column: 7
               Seq primer: T3
               High quality sequence stop: 350
               POLYA=No.

FEATURES       Location/Qualifiers
               source
               1..691
               /organism="Oryza sativa"
               /mol_type="mRNA"
               /cultivar="Pokkali"
               /db_xref="taxon:4530"
               /clone="OG03G07"
               /tissue_type="leaf"
               /dev_stage="7 day-old (treated with 150 mM NaCl for 1, 2
               and 7 days)."
               /lab_host="Escherichia coli"
               /clone_lib="OG"
               /note="Vector: UNI-ZAP XR, pBluescriptSK- (Stratagene);
               Site 1: EcoRI; Site 2: XhoI"
BASE COUNT     105 a 253 c 203 g 130 t
ORIGIN
Query Match    2.0%; Score 50.2; DB 10; Length 691;
Best Local Similarity 50.6%; Pred. No. 0.023;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1958 GGAAGCGTCACGAGCTCAACACGCTTTCGGCCAAATTCGGCAGCGGGGCTCTCAA 2017
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 283 GGTGGCTTCGCCACGGCTACTCTCTCGTGGGGCCCTTCGGCTCACGGGCCGCTCG 342
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2018 GAATGAGGAGATCTTGTGCTCGCGAGCAGCTGGCCACCCGTCTCACAGGTGTCTACC 2077
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 343 CAACAGCGGTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 402
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2078 CGCCACTGCAACTTCGGCGGCGGCGGCTTTCGCTTTCGGCGGAACATGCTGTCTCCGT 2137
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 403 CAGCGTCTGCTTCCATGTACGGGTGCTCTCTTCGGCGAGGCGGCGGCGGCGGCGGCGG 462
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2138 GCTGAAGTCGGAGGCAAGCTCTCCATCATCAAGTCGTTGCGGAGACTGCGGTGGCG 2196
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 463 GCCGAGCTGACGCTCACCGGCGCAAGAGGAGGCGCAAGCTGCAGACCGCGGAGC 521
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
BF430789
LOCUS          BF430789              705 bp mRNA linear EST 20-FEB-2001
DEFINITION    OG04F10T3 OG Oryza sativa cDNA clone OG04F10 5', similar to
               PHOTOSYSTEM I REACTION CENTRE SUBUNIT XI PRECURSOR (SUBUNIT V),
               mRNA sequence.
ACCESSION     BF430789
VERSION       BF430789.1 GI:11442890
KEYWORDS      EST.
SOURCE        Oryza sativa

```

```

ORGANISM       Oryza sativa
SOURCE         Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
               Ehrhartoideae; Oryzeae; Oryza.
REFERENCE      1 (bases 1 to 705)
AUTHORS        Cushman, J.C. and Nomura, M.
TITLE          A cDNA library from Oryza sativa cv. Pokkali leaves enriched for
               salt-responsive genes
JOURNAL        Unpublished
COMMENT        Contact: Cushman JC
               Department of Biochemistry
               University of Nevada
               MS200, Reno, NV 89557-0014, USA
               Tel: 775-784-1918
               Fax: 775-784-1650
               Email: jcushman@unr.edu
               PCR Primers
               FORWARD: T7
               BACKWARD: T3
               Plate: OG04 row: F column: 10
               Seq primer: T3
               High quality sequence stop: 350
               POLYA=No.

FEATURES       Location/Qualifiers
               source
               1..705
               /organism="Oryza sativa"
               /mol_type="mRNA"
               /cultivar="Pokkali"
               /db_xref="taxon:4530"
               /clone="OG04F10"
               /tissue_type="leaf"
               /dev_stage="7 day-old (treated with 150 mM NaCl for 1, 2
               and 7 days)."
               /lab_host="Escherichia coli"
               /clone_lib="OG"
               /note="Vector: UNI-ZAP XR, pBluescriptSK- (Stratagene);
               Site 1: EcoRI; Site 2: XhoI"
BASE COUNT     117 a 253 c 213 g 122 t
ORIGIN
Query Match    2.0%; Score 50.2; DB 10; Length 705;
Best Local Similarity 50.6%; Pred. No. 0.023;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1958 GGAAGCGTCACGAGCTCAACACGCTTTCGGCCAAATTCGGCAGCGGGGCTCTCAA 2017
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 335 GGTGGCTTCGCCACGGCTACTCTCTCGTGGGGCCCTTCGGCTCACGGGCCGCTCG 394
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2018 GAATGAGGAGATCTTGTGCTCGCGAGCAGCTGGCCACCCGTCTCACAGGTGTCTACC 2077
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 395 CAACAGCGGTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 454
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2078 CGCCACTGCAACTTCGGCGGCGGCGGCTTTCGCTTTCGGCGGAACATGCTGTCTCCGT 2137
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 455 CAGCGTCTGCTTCCATGTACGGGTGCTCTCTTCGGCGAGGCGGCGGCGGCGGCGGCGG 514
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2138 GCTGAAGTCGGAGGCAAGCTCTCCATCATCAAGTCGTTGCGGAGACTGCGGTGGCG 2196
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 515 GCCGAGCTGACGCTCACCGGCGCAAGAGGAGGCGCAAGCTGCAGACCGCGGAGC 573
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
CB685121/c
LOCUS          CB685121              710 bp mRNA linear EST 09-APR-2003
DEFINITION    OSJNEF15E15.r OSJNEF Oryza sativa (japonica cultivar-group) cDNA
               clone OSJNEF15E15 3', mRNA sequence.
ACCESSION     CB685121
VERSION       CB685121.1 GI:29688846
KEYWORDS      EST.
SOURCE        Oryza sativa (japonica cultivar-group)
ORGANISM       Oryza sativa (japonica cultivar-group)
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

REFERENCE
AUTHORS Ehrhartoidae; Oryzae; Oryza.
1 (bases 1 to 710)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea

JOURNAL
COMMENT Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 15 row: E column: 15
Seq primer: gga aac agc tat gac cat g.

FEATURES
source
1. .710
Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEF1515"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEF"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control"

BASE COUNT 154 a 201 c 230 g 125 t
ORIGIN

Query Match 2.0%; Score 50.2; DB 14; Length 710;
Best Local Similarity 50.6%; Pred. No. 0.023;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 1958 GGAAGCGCTCACGAACGTCAACAGCCTTTTCGGCCAAATTCGGCAGCGGGCTCTCTCAA 2017
DB 577 GGTGCGCTCGCCACCGCTACCTCTCTGTCGGGCTTTCGGCTCACCGGCGGCTCG 518
QY 2018 GAATGAGGAGATCTCTGCTCGCCAGACCTGGCCACCGCTCTCACAGGTGTCTACCC 2077
DB 517 CAACACGCGGTGCACGCGGAGCGCGCTTCGGCTCGCGGCTCTCTCGCCATCTCT 458
QY 2078 CGCCACTGACAACTTCGCGCGCGCTTTCGCTTCGCGGAAATGCTGTCTCTCGT 2137
DB 457 CAGCGTCTGCTTACCATGTACGGGCTGCGCTTTCGGGAGGGCGAGCGCTCCACGGC 398
QY 2138 GCTGAAGTCGGAGGCAACGCTCTCCATCATCAAGTCCGTTGGCGAGACTGCGCTCGGG 2196
DB 397 GCCGAGCTGACCTCACCGCGCGCAAGAGGAGGCGGCAAGCTGCAGACGCGCGACG 339

RESULT 13
CB685782
LOCUS OSJNEF16102.f OSJNEF Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone OSJNEF16102 5', mRNA sequence.
ACCESSION CB685782
VERSION CB685782.1 GI:29689507
KEYWORDS EST.

SOURCE
ORGANISM Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoidae; Oryzae; Oryza.
REFERENCE
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea

JOURNAL
COMMENT Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 16 row: I column: 02
Seq primer: gta aaa cga cgg cca gtc.

FEATURES
source
1. .776
Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEF16102"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEF"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control"

BASE COUNT 131 a 263 c 232 g 150 t
ORIGIN

Query Match 2.0%; Score 50.2; DB 14; Length 776;
Best Local Similarity 50.6%; Pred. No. 0.023;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 1958 GGAAGCGCTCACGAACGTCAACAGCCTTTTCGGCCAAATTCGGCAGCGGGCTCTCTCAA 2017
DB 364 GGTGCGCTCGCCACCGCTACCTCTCTGTCGGGCTTTCGGCTCACCGGCGGCTCG 423
QY 2018 GAATGAGGAGATCTCTGCTCGCCAGACCTGGCCACCGCTCTCACAGGTGTCTACCC 2077
DB 424 CAACACGCGGTGCACGCGGAGCGCGCTTCGGCTCGCGGCTCTCTCGCCATCTCT 483
QY 2078 CGCCACTGACAACTTCGCGCGCGCTTTCGCTTCGCGGAAATGCTGTCTCTCGT 2137
DB 484 CAGCGTCTGCTTACCATGTACGGGCTGCGCTTTCGGGAGGGCGAGCGCTCCACGGC 543
QY 2138 GCTGAAGTCGGAGGCAACGCTCTCCATCATCAAGTCCGTTGGCGAGACTGCGCTCGGG 2196
DB 544 GCCGAGCTGAGCTCACCGGCGCGCAAGAGGAGGCGGCAAGCTGCAGACGCGGACG 602

RESULT 14
CB651317/c
LOCUS OSJNEB16D17.3', mRNA sequence.

DEFINITION clone OSJNEB16D17 3', mRNA sequence.
ACCESSION CB651317
VERSION CB651317.1 GI:29646310
KEYWORDS EST.

SOURCE
ORGANISM Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoidae; Oryzae; Oryza.

REFERENCE
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL
COMMENT Unpublished
Contact: Rod Wing

Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 16 row: D column: 17
Seq primer: gga aac agc tat gac cat g.

Location/Qualifiers
1..781
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEb16D17"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEb"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 24 hrs after inoculation with Rice Blast (Che 86061)"

BASE COUNT 167 a 216 c 252 g 146 t
ORIGIN

Query Match 2.0%; Score 50.2; DB 14; Length 781;
Best Local Similarity 50.6%; Pred. No. 0.023;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1958 GGAAGGCGTCACGACGCTCAACACGCTTTTCGGCCAAATTCGGCGACGCGGCGCTCTCTCAA 2017
DB 588 GTTCGGCTCGCCACGGCTACTCTCTGTCGGGCCCTTCGGCTCAACGGCCGCTCGG 529

QY 2018 GAATGAGGAGATCTCTGCTCGCGACGACTGCGCACCCCTCTCAAGGTGTCTACCC 2077
DB 528 CAACACGCGGTGCACGCGCGCGCGCTCGCGCGCGCGCGCTCGCTCGCATCT 469

QY 2078 CGCCACTGACACTTTCGGCGCGCGCTTCTGCTTCGGCGACATGCTGCTCTCGT 2137
DB 468 CAGCGTCTGCTTACCATGTACGGGTGCGCTTTCGGCGAGGCGGCGCTCCACGCG 409

QY 2138 GCTGAAGTCGAGGACGCAAGCTCTCTCATCATCAAGTCCGTTGCGGAGACTGCGGCGG 2196
DB 408 GCGGAGCTGAGCTCACCGCGCGCAAGAGAGGCGCGACAGCTGCAGACCGCGACG 350

RESULT 15
CB666855/c
LOCUS
DEFINITION
CB666855 793 bp mRNA linear EST 09-APR-2003
OSJNEb13019.r OSJNEb Oryza sativa (japonica cultivar-group) CDNA
clone OSJNEb13019 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
EST.
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 793)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,R., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 13 row: O column: 19
Seq primer: gga aac agc tat gac cat g.

Location/Qualifiers
1..793
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEb13019"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEb"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"

BASE COUNT 171 a 219 c 252 g 151 t
ORIGIN

Query Match 2.0%; Score 50.2; DB 14; Length 793;
Best Local Similarity 50.6%; Pred. No. 0.023;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1958 GGAAGGCGTCACGACGCTCAACACGCTTTTCGGCCAAATTCGGCGACGCGGCGCTCTCAA 2017
DB 595 GGTGGCGCTCGCCACGGCTACTCTCTGTCGGGCCCTTCGGCTCAACGGCCGCTCGG 536

QY 2018 GAATGAGGAGATCTCTGCTCGCGACGACTGCGCACCCGCTCTCAAGGTGTCTACCC 2077
DB 535 CAACACGCGGTGCACGCGCGCGCGCTTCGGCGCGCGCGCTCTGCTGCTCTCT 476

QY 2078 CGCCACTGACAACTTTCGGCGCGCGGTTTTCGCTTCGCGCGGACATGCTGCTCTCGT 2137
DB 475 CAGCGTCTGCTTACCATGTACGGGTGCGCTTTCGGCGAGGCGGCGCTCCACGCGC 416

QY 2138 GCTGAAGTCGAGGACGCAAGCTCTCTCATCATCAAGTCCGTTGCGGAGACTGCGGCGG 2196
DB 415 GCGGAGCTGAGCTCACCGCGCGCAAGAGAGGCGCGACAGTTCAGACCGCGCGACG 357

Search completed: January 12, 2004, 22:32:16
Job time : 5111 secs

THIS PAGE BLANK (USPTO)